



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA INDUSTRIAL

ANÁLISIS DE USUARIOS EN APLICACIÓN MÓVIL Y
ECONOMÍA DEL COMPORTAMIENTO EN REPORTE DE USUARIOS
PARA MEJORAR EL TRANSPORTE PÚBLICO

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO CIVIL INDUSTRIAL

CLAUDIO ANDRÉS MENA CAVIERES

PROFESOR GUÍA:
DANIEL SCHWARTZ PERLROTH

PROFESORA CO-GUÍA:
MARCELA MUNIZAGA MUÑOZ

MIEMBRO DE LA COMISIÓN:
JACQUELINE ARRIAGADA FERNÁNDEZ

SANTIAGO DE CHILE

2020

RESUMEN DE LA MEMORIA PARA OPTAR AL
TÍTULO DE INGENIERO CIVIL INDUSTRIAL
POR: CLAUDIO ANDRÉS MENA CAVIERES
FECHA: 24/FEBRERO/2020
PROF. GUÍA: SR. DANIEL SCHWARTZ PERLROTH

ANÁLISIS DE USUARIOS EN APLICACIÓN MÓVIL Y ECONOMÍA DEL
COMPORTAMIENTO EN REPORTES DE USUARIOS PARA MEJORAR EL
TRANSPORTE PÚBLICO

Transapp es una aplicación colaborativa móvil para *smartphone* fundada el 2016, en donde los usuarios pueden consultar tiempos de llegada de buses, emitir reportes acerca del estado de la infraestructura (ya sea buena o mala) y seguir sus viajes en bus vía la aplicación para luego evaluar su recorrido. A partir de lo anterior se define como un sistema inteligente de transporte (ITS), el cual no sólo utiliza información de las autoridades de transporte para informar a los usuarios, sino que también utiliza la información de otras personas, método conocido como *crowdsourcing*. Lamentablemente, hay evidencia en diferentes plataformas de problemas de participación en donde una baja proporción de usuarios aporta contenido.

De manera de velar por la sustentabilidad de la aplicación, se ha propuesto comprender el comportamiento de los usuarios. En particular, se desarrolla un análisis de datos de las herramientas de consultas, reportes y viajes, cuyas principales aristas son: primero, un análisis descriptivo; segundo, un modelo explicativo de la evaluación entregada por los usuarios, donde se demuestra que los usuarios que más realizan esta actividad evalúan más positivamente sus viajes, revelando un sesgo de selección; tercero, un análisis de sobrevivencia para modelar la inactividad, en el cual se muestra que quienes participan de las herramientas tempranamente tienen menor riesgo de volverse inactivos; y cuarto, un estudio de los reportes, donde se concluye que *Transapp* sí sufre del problema de participación.

El presente trabajo de investigación, además, examina qué incentivos aumentan la proporción de usuarios que generan contenido en una plataforma *online*, a través de un experimento de campo en *Transapp*. Para esto, se envía aleatoriamente a los usuarios 3 notificaciones motivándoles a reportar mediante diferentes incentivos: prosocial (“reportar ayuda al resto”), económico monetario (“participa por carga bip!”) o una combinación de ambos. Dado que los usuarios se autoseleccionan para recibir el tratamiento intensivo (abrir la notificación), se analiza el efecto de la recepción de la notificación (libre de endogeneidad), la cual muestra una fuerte heterogeneidad dependiendo del uso previo de los usuarios. En aquellos que no reportaban o lo hacían ocasionalmente, los incentivos con componente económico son los más efectivos. En cambio, en aquellos que ya reportaban de manera frecuente, si bien ambos incentivos son positivos, se recomienda el prosocial ya que es más coste-efectivo. Para futuros estudios se propone la realización de experimentos adicionales con un mayor tamaño de muestra. Finalmente, este trabajo permite desarrollar nuevas herramientas que potencian: la capacidad que tienen los usuarios de ser agentes activos en la mejora del transporte público mediante demandas sobre mejor calidad, y los beneficios de compartir información.

*Dedicada a mi
querida familia
por cada esfuerzo
que me permite
escribir esto*

*“La pobreza no es natural, es creada por el humano y
puede superarse mediante acciones de los seres humanos.
Y erradicar la pobreza no es un acto de caridad,
es un acto de justicia”
Nelson Mandela*

*Piano piano,
si va lontano...*

Agradecimientos

Este documento marca el final de una etapa larga, compleja, desafiante y que me hizo crecer en formas que jamás esperaba. Decir gracias a cada una de las personas que me ha acompañado es difícil, pero es una oportunidad única para alguien que no suele hacerlo.

Quisiera primero agradecer a las personas más importantes en mi vida. Mi familia. Les agradezco por cada día que se esforzaron por permitirme crecer en un ambiente acogedor, cálido y seguro. Gracias Jennifer por tu hermandad eterna, por cuidarme hasta el día de hoy, por las aventuras y tonterías. Gracias Papá y Mamá por criarme sin preocupaciones, por brindarme la mejor educación posible, por siempre potenciarme al siguiente nivel, por su cariño y amor. Quisiera destacar a mis abuelas/os, Yeye, Remigio, Margarita y Señora Ana, cuyo esfuerzo y valor diario está presente, incluso en las palabras de este documento. Gracias a mis perros Olaf Kowalski, Kiara, Doky y Snoopy, por alegrarme los días y llegadas a casa cuando venía agotado.

Gracias a mis amigos/as que me acompañaron durante la Universidad. Gracias a Dani, Fran, Vicho y al Franco, por hacer el plan común menos terrible, por los almuerzos entretenidos y por las noches de juegos cada vez más complejos de entender. Gracias a quienes me acompañaron en Industrias, en especial a Richi, por tus vitales consejos y por ser el mejor oyente, tu sabiduría aún me invita a oír más y decir menos. Gracias Gabi, desde el preuniversitario, por tu amistad.

Gracias al equipo de Tutoría 2017 y 2018, hoy en día Tutoresantes, por la confianza, buena onda, por entretener el final del camino y por haberme permitido crecer junto a ustedes. Quisiera agradecer también a las personas de las oficinas de Docencia y Gestión de Carrera, ya que fue un agrado trabajar y compartir con ustedes. Especialmente a Gaby por su paciencia y a Mónica por su ayuda y guía en este proceso de titulación. De igual forma, gracias a Helia, quien se preocupó diariamente de que pudiera trabajar en un espacio limpio, la conversación matutina y los buenos deseos; también a Mirtha, quien me ayudó más de una vez con algún libro y por sus buenas conversaciones; y a Mariana, quien me acompañó en el tortuoso viaje de la Estación Central a la Universidad motivándome a trabajar.

Agradezco al profesor Daniel por haber confiado en mí, fue un gusto haber trabajado con usted en la docencia en Aplica. De igual manera, agradezco a la profesora Marcela quien me desafía a presentar estos resultados en Portugal. Agradezco a ambos su guía, la infinita paciencia, los consejos, y por haberme permitido haber trabajado con ustedes, el cual es un gran honor para mí.

Gracias al equipo de *Transapp* por permitirme realizar este trabajo con ustedes y ser parte de su equipo. En especial a Mauricio, por la paciencia y ayuda con los datos; a Felipe, por programar mi experimento y permitirme trabajar en su aplicación; y Jacqueline por su ayuda transversal durante el proceso. Es motivador ver gente comprometida con el Transporte Público.

Gracias a quienes me educaron. Gracias a las profesoras de la Escuela de Lenguaje que me enseñaron a hablar; a las profesoras/os de mis colegios y el Instituto Nacional, en especial a Danissa y Nancy, quienes me enseñaron a reflexionar, vieron potencial en mí y me motivaron a seguir aprendiendo. Gracias a la Universidad de Chile por formarme como ingeniero.

Si bien esto marca el final de una etapa, estoy muy emocionado por lo que ya vendrá.

Muchas gracias.

Tabla de contenido

| | | |
|------------|---|----|
| Capítulo 1 | Introducción..... | 1 |
| Capítulo 2 | Revisión bibliográfica..... | 5 |
| 2.1 | Sistemas inteligentes de transportes (ITS) | 5 |
| 2.2 | Participación en plataformas online..... | 7 |
| 2.3 | Motivaciones para participar en plataformas online | 8 |
| Capítulo 3 | Análisis de datos..... | 10 |
| 3.1. | Usuarios de la aplicación según consultas de paraderos..... | 11 |
| 3.2. | Análisis de consultas de paraderos..... | 14 |
| 3.2.1 | Análisis descriptivo | 14 |
| 3.2.2 | Análisis local | 18 |
| 3.3. | Análisis de reportes..... | 21 |
| 3.3.1 | Análisis descriptivo | 22 |
| 3.3.2 | Análisis local | 25 |
| 3.3.3 | Problema de participación | 33 |
| 3.4. | Análisis de viajes y evaluaciones..... | 35 |
| 3.4.1 | Análisis descriptivo | 35 |
| 3.4.2 | Análisis de viajes y evaluaciones locales | 40 |
| 3.4.3 | Análisis descriptivo de la nota del viaje..... | 44 |
| 3.4.4 | Modelo para explicar variaciones en la nota del viaje..... | 48 |
| 3.5 | Conclusiones del trabajo de datos..... | 56 |
| Capítulo 4 | Análisis de sobrevivencia de usuarios | 58 |
| 4.1 | Análisis según actividades de la semana 1 | 60 |
| 4.2 | Análisis según actividades en el período | 63 |
| 4.3 | Conclusiones del análisis de sobrevivencia | 66 |
| Capítulo 5 | Experimento de campo | 67 |
| 5.1 | Metodología | 67 |
| 5.2 | Condiciones experimentales | 68 |
| 5.3 | Descripción de los participantes..... | 70 |
| 5.4 | Exposición al tratamiento..... | 72 |

| | | |
|------------|---|-----|
| 5.5 | Análisis de resultados | 75 |
| 5.5.1 | Análisis de resultados – según recepción..... | 75 |
| 5.5.2 | Análisis de resultados – según apertura..... | 81 |
| 5.6 | Conclusiones del experimento de campo..... | 89 |
| 5.6.1 | Conclusiones según recepción..... | 89 |
| 5.6.2 | Conclusiones según apertura..... | 90 |
| 5.6.3 | Conclusiones teóricas | 91 |
| 5.7 | Discusión y futuros estudios | 93 |
| Capítulo 6 | Conclusiones generales..... | 94 |
| | Bibliografía | 95 |
| | Anexos..... | 99 |
| 1. | Apéndice A: Estadísticas de consultas de paradero por semana..... | 99 |
| 2. | Apéndice B: Estadísticas de reportes por semana..... | 101 |
| 3. | Apéndice C: Estadísticas de reportes acerca de buses o paradero, por semana. | 103 |
| 4. | Apéndice D: Estadísticas de reportes positivos y pesimistas, según semana. | 105 |
| 5. | Apéndice E: Estadísticas de reportes confirmativos y negativos, según semana..... | 107 |
| 6. | Apéndice F: Estadísticas de viajes y evaluaciones semanales | 109 |
| 7. | Apéndice G: Estadísticas de recorridos..... | 112 |
| 8. | Apéndice H: Modelo autoregresivo de los recorridos viajados y no evaluados..... | 113 |
| 9. | Apéndice I: Notas de la base de datos | 116 |
| 10. | Apéndice J: Velocidad del viaje según día de la semana e intervalo horario | 117 |
| 11. | Apéndice K: Condiciones experimentales..... | 119 |
| 12. | Apéndice L: Análisis de resultados de experimento considerando usuarios que no reciben las notificaciones | 121 |
| 13. | Apéndice M: Análisis de resultados omitiendo segundo día..... | 122 |
| 13.1 | Resultados según recepción..... | 122 |
| 13.2 | Resultados según apertura | 124 |
| 14. | Apéndice N: Resultados de regresión lineal 2SLS – para estimación CACE modelo tipo de usuario | 125 |
| 14.1 | Grupo base “Reportador Tímido” | 125 |
| 14.2 | Grupo base “Reportador” | 126 |

Índice de tablas

| | |
|---|----|
| Tabla 1: Descripción de las bases de datos de Transapp a analizar. | 10 |
| Tabla 2: Glosario de términos frecuentes asociados a usuarios. | 10 |
| Tabla 3: Cantidad mensual de usuarios activos, nuevos, “de baja” y la variación mensual. | 11 |
| Tabla 4: Cantidad y proporción de usuarios activos en determinadas semanas, considerando toda la muestra. | 12 |
| Tabla 5: Cantidad y proporción de usuarios activos en determinadas semanas, considerando la muestra factible. | 13 |
| Tabla 6: Estadísticas descriptivas de las consultas de paraderos mensuales. | 14 |
| Tabla 7: Cantidad y proporción de usuarios que realizan consultas únicas en todo el período. | 17 |
| Tabla 8: Cantidad de consultas únicas de paraderos, usuarios consultores y promedio de consultas únicas por usuario según intervalo horario del día. | 18 |
| Tabla 9: Consultas únicas por paradero y promedio de consultas únicas por usuario, según día de la semana. | 19 |
| Tabla 10: Filtro de la base de datos de reportes. | 21 |
| Tabla 11: Estadísticas descriptivas de los reportes y reportadores, mensualmente. | 22 |
| Tabla 12: Cantidad y proporción de usuarios que realizan viajes en determinadas semanas. | 24 |
| Tabla 13: Cantidad de reportes mensuales, acerca de paradero y bus, mensualmente. | 25 |
| Tabla 14: Clasificación de reportes positivos y pesimistas en reportes de buses y paraderos. | 27 |
| Tabla 15: Cantidad de reportes mensuales, acerca de paradero y bus. | 28 |
| Tabla 16: Cantidad de reportes confirmativos y negativos, semanalmente. | 29 |
| Tabla 17: Cantidad de reportes y usuarios reportadores, según intervalo horario. | 31 |
| Tabla 18: Cantidad de reportes y usuarios reportadores, según día de la semana. | 32 |
| Tabla 19: Contraste de reportes en Transapp con la Regla de Nielsen. | 33 |
| Tabla 20: Estadísticas asociadas a la Regla de Nielsen en Transapp. Cantidad de reportes, reportadores y proporciones de reportes según la clasificación de usuario. | 34 |
| Tabla 21: Cantidad de viajes, evaluaciones y usuarios viajeros y evaluadores, mensualmente. | 36 |
| Tabla 22: Cantidad y proporción de usuarios que realizan viajes en determinadas semanas. | 38 |
| Tabla 23: Cantidad y proporción de usuarios que realizan evaluaciones de viajes en determinadas semanas. | 39 |
| Tabla 24: Frecuencia de viajes y evaluaciones realizadas según hora del día. | 40 |
| Tabla 25: Resultados regresión de proporción de viajes evaluados según día de la semana. | 41 |
| Tabla 26: Frecuencia de viajes y evaluaciones realizadas según día de la semana. | 42 |
| Tabla 27: Resultados regresión de proporción de viajes evaluados según día de la semana. | 43 |
| Tabla 28: Cantidad de evaluaciones y promedio de la nota según velocidad del viaje evaluado. | 47 |
| Tabla 29: Cantidad de evaluaciones y estadísticas según formato de evaluación. | 48 |
| Tabla 30: Estadísticas de las evaluaciones de formato de 3 evaluaciones. | 49 |
| Tabla 31: Distribución de usuarios según viajes realizados en período 1. | 49 |
| Tabla 32: Resultados estimación modelo probit ordenado de evaluación para 3 evaluaciones. | 50 |

| | |
|---|----|
| Tabla 33: Probabilidades predichas por el modelo de 3 evaluaciones y el de los datos..... | 50 |
| Tabla 34: Resultados efectos marginales de probit ordenado de 3 evaluaciones..... | 51 |
| Tabla 35: Estadísticas de las evaluaciones de formato de 5 evaluaciones..... | 52 |
| Tabla 36: Distribución de usuarios según viajes realizados en período 1..... | 52 |
| Tabla 37: Resultados estimación modelo probit ordenado de evaluación para 5 evaluaciones. | 53 |
| Tabla 38: Probabilidades predichas por el modelo de 5 evaluaciones y el de los datos. | 53 |
| Tabla 39: Resultados efectos marginales de probit ordenado de 5 evaluaciones..... | 54 |
| Tabla 40: Resumen estadísticas de modelación tiempo de sobrevivencia. | 58 |
| Tabla 41: Clasificación de usuarios según acciones durante la primera semana de actividad de consulta de paraderos. | 60 |
| Tabla 42: Resumen estadísticas de modelación tiempo de sobrevivencia, según categoría de usuario en base a sus acciones durante la primera semana de actividad de consulta de paraderos. | 61 |
| Tabla 43: Test log-rank de la igualdad de las funciones de sobrevivencia según categoría de usuarios en base a acciones realizadas en la primera semana de consulta de paraderos. | 62 |
| Tabla 44: Resultados regresión Cox. | 62 |
| Tabla 45: Clasificación de usuarios según acciones durante su período de actividad. | 63 |
| Tabla 46: Resumen estadísticas de modelación tiempo de sobrevivencia, según categoría de usuario según acciones durante su período de actividad. | 63 |
| Tabla 47: Test log-rank de la igualdad de las funciones de sobrevivencia según categoría de usuarios según acciones durante su período de actividad. | 64 |
| Tabla 48: Resultados regresión Cox. | 65 |
| Tabla 49: Configuración y detalle de la notificación de cada condición experimental del día 1...69 | 69 |
| Tabla 50: Criterio de clasificación de usuarios activos para realización de experimento. | 70 |
| Tabla 51: Estadísticas de clasificación de usuarios según reportes realizados. | 70 |
| Tabla 52: Distribución de usuarios en las condiciones experimentales según clasificación de uso en la aplicación..... | 71 |
| Tabla 53: Distribución de usuarios en las diferentes condiciones experimentales..... | 71 |
| Tabla 54: Estadísticas sobre recepción y exposición a los tratamientos de las condiciones experimentales. | 72 |
| Tabla 55: Resultado regresión logit, estudio de diferencias de tasa de recepción entre notificaciones. | 73 |
| Tabla 56: Comparación por pares entre notificaciones para identificar diferencias en tasa de recepción..... | 73 |
| Tabla 57: Estadísticas para estimación del ITT - modelo simple..... | 76 |
| Tabla 58: Resultados modelo logit - simple, para estimación del efecto ITT..... | 77 |
| Tabla 59: Resultados análisis comparación por pares entre notificaciones, del modelo simple en efecto ITT..... | 77 |
| Tabla 60: Resultados modelo logit - tipo de usuario, para estimación del efecto ITT. | 79 |
| Tabla 61: Resultados de efecto ITT a partir de modelo - tipo de usuario..... | 79 |
| Tabla 62: Resultados análisis comparación por pares entre notificaciones, del modelo tipo de usuario en efecto ITT | 80 |
| Tabla 63: Análisis modelo logit para identificar diferencias estadísticas entre tasas de apertura.81 | 81 |
| Tabla 64: Estadísticas para estimación del ITT - modelo simple..... | 82 |

| | |
|--|-----|
| Tabla 65: Estadísticas para estimación del CACE - modelo simple..... | 82 |
| Tabla 66: Resultados modelo lineal - simple, para estimación del efecto CACE..... | 84 |
| Tabla 67: Resultados análisis comparación por pares entre notificaciones, del modelo simple en efecto CACE..... | 84 |
| Tabla 68: Estadísticas para estimación del ITT - modelo tipo de usuario. | 85 |
| Tabla 69: Estadísticas para estimación del CACE - modelo tipo de usuario. | 85 |
| Tabla 70: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, con grupo base control del tipo de usuario “No Reportador”..... | 87 |
| Tabla 71: Resultados de efecto CACE a partir de modelo – tipo de usuario, y contraste con el teórico..... | 87 |
| Tabla 72: Resultados análisis comparación por pares entre notificaciones, del modelo tipo de usuario en efecto CACE..... | 88 |
| Tabla 73: Estadísticas descriptivas de las consultas de paraderos mensuales..... | 99 |
| Tabla 74: Estadísticas descriptivas de reportes y reportadores, semanalmente..... | 101 |
| Tabla 75: Estadísticas semanales de la cantidad de reportes de buses y paraderos, y su proporción. | 103 |
| Tabla 76: Estadísticas semanales de la cantidad de reportes positivos y pesimistas. | 105 |
| Tabla 77: Estadísticas semanales de la cantidad de reportes confirmativos y negativos, y su proporción. | 107 |
| Tabla 78: Cantidad de viajes, evaluaciones y usuarios viajeros y evaluadores, semanalmente... | 109 |
| Tabla 79: Resultados estimación modelo autoregresivo de cuatro lags de la cantidad de recorridos viajados y no evaluados, semanalmente. | 113 |
| Tabla 80: Resultados estimación modelo simple de recorridos no evaluados..... | 114 |
| Tabla 81: Resultados test Breusch-Godfrey para modelo simple..... | 114 |
| Tabla 82: Resultados modelo autoregresivo de dos lags. | 115 |
| Tabla 83: Resultados test Breusch-Godfrey para modelo autoregresivo de dos lags. | 115 |
| Tabla 84: Resultados estimación modelo autoregresivo de cuatro lags de la cantidad de recorridos viajados y no evaluados, semanalmente. | 115 |
| Tabla 85: Resultados test Breusch-Godfrey del modelo autoregresivo de 4 lags. | 115 |
| Tabla 86: Valores de evaluaciones en el informe según base de datos..... | 116 |
| Tabla 87: Resultados regresión velocidad del bus según intervalo horario. | 117 |
| Tabla 88: Resultados regresión velocidad del bus según día de la semana. | 118 |
| Tabla 89: Resultados generales de estimación de velocidad de bus según intervalo horario y día de la semana..... | 118 |
| Tabla 90: Configuración y detalle de la notificación de cada condición experimental del día 2. | 119 |
| Tabla 91: Configuración y detalle de la notificación de cada condición experimental del día 3. | 120 |
| Tabla 92: Proporción de usuarios no receptores de notificaciones que reportaron, según condición experimental. | 121 |
| Tabla 93: Resultados estimación modelo logit en usuarios no receptores de notificaciones..... | 121 |
| Tabla 94: Resultados modelo logit - simple, para estimación del efecto ITT omitiendo datos del segundo día..... | 122 |
| Tabla 95: Resultados modelo logit - tipo de usuario, para estimación del efecto ITT, omitiendo datos del segundo día. | 123 |

| | |
|--|-----|
| Tabla 96: Resultados de efecto ITT a partir de modelo - tipo de usuario, omitiendo datos del segundo día..... | 123 |
| Tabla 97: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, omitiendo datos del segundo día..... | 124 |
| Tabla 98: Resultados de efecto CACE a partir de modelo - simple, omitiendo datos del segundo día..... | 124 |
| Tabla 99: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, con grupo base control del tipo de usuario “Reportador Tímido”..... | 125 |
| Tabla 100: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, con grupo base control del tipo de usuario “Reportado”..... | 126 |

Índice de ilustraciones

| | |
|--|----|
| Ilustración 1: Principales herramientas de la aplicación Transapp..... | 2 |
| Ilustración 2: Flujo de información potencial de Transapp. | 3 |
| Ilustración 3: Explicación de usuario activo e inactivo en la aplicación, en base a la consulta de paraderos. | 11 |
| Ilustración 4: Distribución de usuarios según número de semanas de actividad..... | 12 |
| Ilustración 5: Distribución semanal de las consultas realizadas y las consultas únicas..... | 15 |
| Ilustración 6: Número de usuarios activos y usuarios consultores, semanalmente..... | 16 |
| Ilustración 7: Promedio de consultas por usuario consultor, semanalmente..... | 16 |
| Ilustración 8: Histograma de usuarios según consultas únicas (izquierda) y logaritmo de consultas únicas (derecha). | 17 |
| Ilustración 9: Consultas únicas por paradero y promedio de consultas únicas por usuario, según hora del día..... | 19 |
| Ilustración 10: Consultas únicas de paradero y promedio de consultas únicas por usuario, según día de la semana..... | 20 |
| Ilustración 11: Cantidad de reportes semanales, usuarios reportadores y proporción de reportadores, semanalmente. | 23 |
| Ilustración 12: Promedio de reportes por usuario, semanalmente. | 23 |
| Ilustración 13: Proporción de usuarios según reportes realizados (izquierda) y logaritmo de reportes realizados (derecha)..... | 24 |
| Ilustración 14: Reportes de paraderos y buses semanalmente. | 26 |
| Ilustración 15: Proporción de reportes de buses y paraderos respecto al total realizado, semanalmente. | 26 |
| Ilustración 16: Reportes semanales positivos y pesimistas..... | 28 |
| Ilustración 17: Proporción de reportes positivos y pesimistas respecto al total realizado, semanalmente. | 29 |
| Ilustración 18: Reportes confirmativos y negativos semanales. | 30 |
| Ilustración 19: Proporción semanal de reportes confirmativos y negativos respecto al total realizado. | 30 |
| Ilustración 20: Reportes realizados y promedio de reportes por usuario según intervalo horario. | 31 |
| Ilustración 21: Reportes realizados y promedio de reportes por usuario según día de la semana. | 32 |
| Ilustración 22: Cantidad de viajes, evaluaciones realizadas, usuarios viajeros y usuarios evaluadores, semanalmente. | 37 |
| Ilustración 23: Cantidad promedio de viajes y evaluaciones por usuario, semanalmente. | 37 |
| Ilustración 24: Proporción de viajes evaluados y proporción de viajeros que evalúan viajes, según semana..... | 38 |
| Ilustración 25: Proporción de usuarios según viajes realizados (izquierda), y según evaluaciones realizadas (derecha)..... | 39 |
| Ilustración 26: Consultas, evaluaciones y proporción de viajes evaluados, según hora del día. | 41 |

| | |
|---|-----|
| Ilustración 27: Consultas, evaluaciones y proporción de viajes evaluados, según día de la semana. | 43 |
| Ilustración 28: Histogramas de evaluación promedio, diferenciando usuarios según cantidad de viajes evaluados. | 45 |
| Ilustración 29: Variación semanal de la evaluación promedio de los viajes evaluados. | 45 |
| Ilustración 30: Variación de la evaluación promedio de los viajes evaluados según intervalo horario. | 46 |
| Ilustración 31: Variación de la nota promedio de los viajes evaluados según día de la semana. .. | 47 |
| Ilustración 32: Variación de la nota promedio de las evaluación según velocidad del bus. | 48 |
| Ilustración 33: Probabilidades de sobrevivencia (izquierda) y de falla (derecha) de Kaplan-Meier considerando todos los usuarios que consultan tiempos de espera en paraderos en el período. | 59 |
| Ilustración 34: Riesgo suavizado (izquierda) y riesgo acumulado de Nelson-Aalen (derecha) de los usuarios consultores según semanas de actividad. | 59 |
| Ilustración 35: Probabilidad de sobrevivencia de Kaplan-Meier según categoría de usuarios en base a actividades realizadas en primera semana de consulta de paraderos. | 61 |
| Ilustración 36: Probabilidad de sobrevivencia de Kaplan-Meier según categoría de usuarios en base a acciones durante su período de actividad. | 64 |
| Ilustración 37: Riesgo suavizado de las categorías de usuario según las actividades realizadas durante su período de actividad. | 65 |
| Ilustración 38: Ejemplo de recepción de todas las notificación push enviadas y la apertura de un mensaje en la bandeja de entrada de los usuarios. | 67 |
| Ilustración 39: Resultados resumidos del efecto ITT - Recepción en la probabilidad de emitir reporte. | 89 |
| Ilustración 40: Resultados teóricos resumidos del efecto CACE - Apertura en la probabilidad de emitir reporte. | 90 |
| Ilustración 41: Recorridos semanales de buses que han sido viajados y evaluados por los usuarios. | 112 |
| Ilustración 42: Cantidad de recorridos viajados, pero no evaluados, semanalmente. | 112 |
| Ilustración 43: Autocorrelación de los residuos del modelo simple. | 114 |
| Ilustración 44: Residuos de modelo autoregresivo de 4 lags. | 116 |

Capítulo 1

Introducción

Gracias al avance tecnológico actual, han surgido una gran cantidad de aplicaciones móviles para teléfonos inteligentes que otorgan información útil para los usuarios (Brazil & Caulfield, 2013) para la planificación de sus viajes en transporte público. En Santiago de Chile estas también representan canales, no operados por el Directorio de Transporte Público Metropolitano (DTPM¹), que ingresan requerimientos, tales como: *Google Maps*, *Moovit*, *Bus Checker*, *Transapp*, entre otras (DTPM, 2019).

En particular, *Transapp* es una “aplicación móvil para teléfonos inteligentes² que otorga ayuda a usuarios del sistema Red Metropolitana de Movilidad³ a informarse del estado del transporte público, de manera simple y dinámica” (Transapp, 2018) lanzada en 2016, que se define como colaborativa (Universidad de Chile, 2016) al permitir que los usuarios *ayuden* al resto. Esto se observa en las diferentes funcionalidades de esta aplicación⁴, en donde los usuarios pueden:

- Solicitar una estimación de tiempo de viaje (mediante un planificador).
- Realizar reportes de los buses acerca de:
 - El estado físico del bus, como las: puertas, ventanas, limpieza, entre otros.
 - La probidad en la conducción del chofer.
 - Ocurrencias de asaltos/hurtos a bordo y/o desórdenes.
 - Saturación y hacinamiento del recorrido, especificando si es posible viajar sentado.
- Evaluar el viaje del bus al terminar el recorrido.
- Realizar reportes de paraderos acerca de:
 - La infraestructura pública como iluminación en mal estado, inexistencia de asiento, techumbre o información de recorridos.
 - Sucesos que provocan inseguridad, como personas en estado de ebriedad, robos y/o desórdenes.
 - Si buses, vacíos o saturados, se detuvieron en paraderos donde el pasajero lo solicitó.
 - La regularidad de la frecuencia de los buses, por ejemplo, si hay buses del mismo recorrido juntos entre sí.

¹ El Directorio de Transporte Público Metropolitano (DTPM) articula y coordina las entidades privadas y públicas, con el objetivo de gestionar el transporte público de Santiago de Chile (DTPM, 2019).

² Disponible para usuarios de teléfonos con sistema operativo Android o iOS.

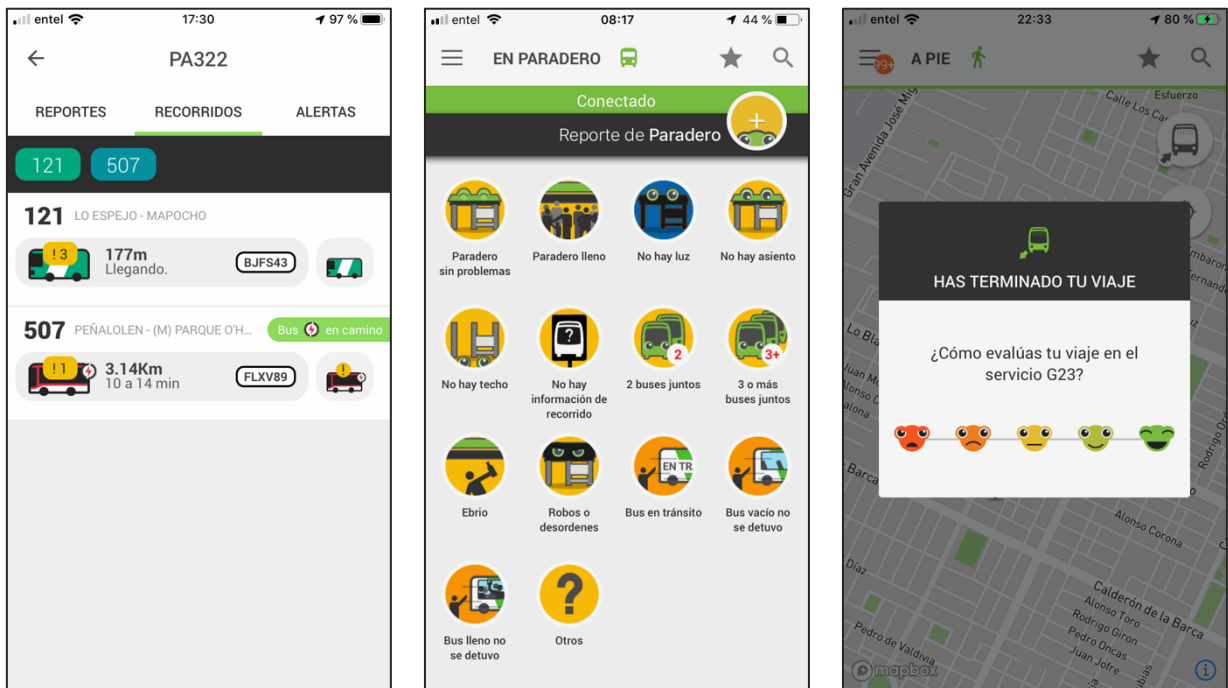
³ Red es el sistema integrado de transporte público de Santiago, que reemplazó gradualmente al Transantiago en 2019 (CNN Chile, 2019).

⁴ Varía dependiendo si el sistema operativo del celular y la versión de la aplicación.

- Transversalmente los usuarios pueden indicar si los reportes de otras personas (que son anónimos) son verdaderos o falso. Por ejemplo, si alguien reporta que el paradero está sin luz, las personas pueden indicar si esto es correcto o incorrecto.
- Recibir notificaciones (mensajes) que se almacenan en la bandeja de entrada. Las cuales se utilizan como envío de alertas en caso de problemas en el transporte de la ciudad por el equipo de *Transapp*.
- Participar de la comunidad de *Transapp* mediante la *gamificación* en donde los usuarios al reportar y viajar en buses acumulan puntos, subiendo así de clasificación y ranking (Antoine, 2018).

La primera imagen de la Ilustración 1 representa la interfaz obtenida al momento de consultar el tiempo de llegada de buses en algún paradero, la segunda corresponde a las posibilidades de reportes que pueden emitir los usuarios sobre un paradero, y finalmente, la tercera corresponde a la evaluación que los usuarios son invitados a responder una vez que su viaje termina.

Ilustración 1: Principales herramientas de la aplicación *Transapp*.



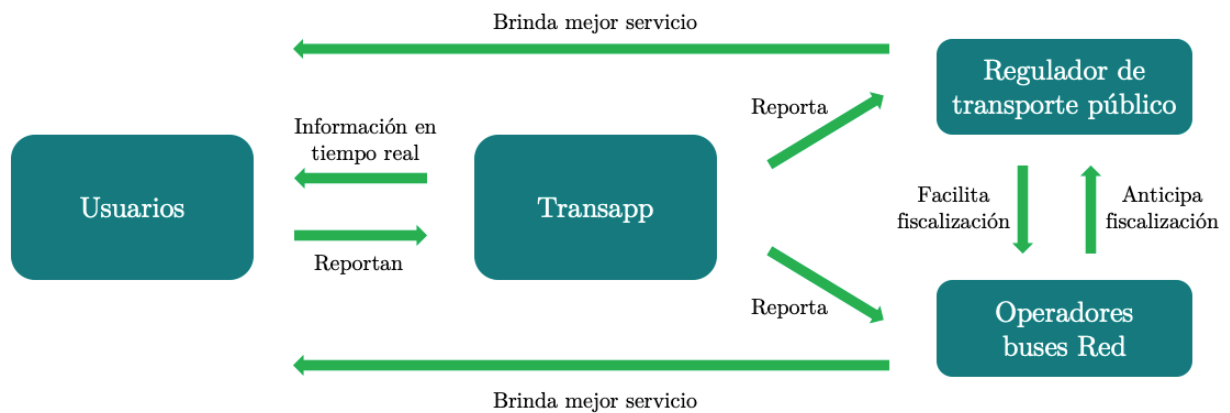
Fuente: obtenidas de la aplicación *Transapp* (2020).

Según la encuesta de satisfacción de empresas del -antiguo- Transantiago que está a cargo del DTPM (Brújula, 2017), cerca del 10% de los encuestados declaró que el mal estado de los buses es una razón para declarar insatisfacción. Así también “poca seguridad / robos” con un 3.9%, “buses sucios” el 3.2%, “choferes imprudentes” con 2.1%; los cuales son todos reportes predeterminados que *Transapp* busca que los usuarios informen. Esta determinación de reportes no es casualidad, ya que requirió de revisión de literatura y análisis de *focus group* por parte del equipo desarrollador previo al lanzamiento de la aplicación.

Una plataforma *online*, según la OCDE (2019) “es un servicio digital que facilita la interacción entre dos o más usuarios interdependientes distintos (ya sean empresas o personas) quienes interactúan a través del servicio vía internet”, por lo que, a partir de las funcionalidades previas, se puede definir a la aplicación *Transapp* como una.

La información levantada por los usuarios es canalizada por *Transapp* y podría, eventualmente, comercializarla con instituciones públicas y/o operadores del sistema Red de buses⁵, de forma tal que éstos puedan brindar un mejor servicio y anticiparse o facilitar eventuales procesos de fiscalización. De igual manera, estos reportes también son valiosos y de ayuda para los otros usuarios de la comunidad quienes puedan realizar una decisión de viaje óptima con mayor información. Esto se puede observar en la Ilustración 2, que muestra el flujo de información potencial y la utilidad de éstos a las diferentes entidades.

Ilustración 2: Flujo de información potencial de *Transapp*.



Fuente: elaboración propia

Siguiendo el esquema anterior, para que *Transapp* sea una aplicación efectiva y útil a los usuarios, debe cerciorarse con contar una cantidad de usuarios que participen activamente reportando y evaluando sus viajes. Ya que así al contar con más información se produce un círculo positivo, al poder ayudar a más usuarios y entregar más información a las organizaciones del sistema.

No obstante, la participación de usuarios en la herramienta de reportes es baja, ya que sólo el 15,73% de los usuarios mensuales⁶ reporta al menos una vez (se profundiza en el capítulo 3). Este problema de bajo involucramiento es frecuente en diferentes plataformas *online*, el cual es resumido en la regla de Nielsen (2006) que indica la existencia de la proporción 90-9-1, en la cual el 90% de los usuarios aporta contenido en la aplicación/plataforma *online*, el 9% lo hace pero ocasionalmente, y el 1% restante lo hace de manera muy frecuente y sus aportes concentran gran magnitud de toda

⁵ Modelo de negocios, hacia las empresas e instituciones del sistema de transporte, en validación.

⁶ Porcentaje promedio mensual de los 14 meses en estudio.

la información levantada por los usuarios, por lo que en el Capítulo 2: Revisión bibliográfica se profundiza sobre la literatura presente.

Dado lo anterior, se propone estudiar cómo incentivar a los usuarios a participar más y cuál incentivo es el más adecuado dependiendo de su actividad histórica, en el levantamiento de información de transporte público en contexto de una aplicación de *crowdsourcing* (se profundiza en la revisión de la literatura). Para responder estas preguntas, se trabaja en un caso de estudio de la aplicación *Transapp* mediante la siguiente estructura.

En el Capítulo 3: Análisis de datos, se analiza las diferentes bases de datos de la aplicación con el objetivo de desplegar información relacionada a las consultas de paraderos, reportes y viajes (y sus evaluaciones) de una manera descriptiva. Tal como se indica anteriormente, se profundiza el problema de participación en la aplicación y se realiza un modelo que explica variaciones en las evaluaciones entregadas por los usuarios de sus viajes.

En el Capítulo 4: Análisis de sobrevivencia de usuarios, se realiza un análisis de usuarios, con el objetivo de identificar diferencias en el uso de la aplicación en la probabilidad de dejar utilizar la aplicación, es decir, que un usuario activo se convierta en inactivo. Se espera que esta información, sumada a la del tercer capítulo, pueda identificar oportunidades de mejora para así facilitar las gestiones que permita potenciar la aplicación, de forma tal de aumentar su competitividad y atractivo respecto a las otras del mercado.

En el Capítulo 5: Experimento de campo, se detalla la metodología, resultados y discusión del experimento de campo aleatorio realizado en la aplicación, el cual permite estudiar y responder a las preguntas planteadas en la investigación de manera exitosa pese a ciertas limitaciones.

Finalmente, se enuncia las conclusiones generales del trabajo realizado en el Capítulo 6: Conclusiones generales.

Capítulo 2

Revisión bibliográfica

En esta sección se profundizará, no extensivamente, sobre los sistemas inteligentes de transportes y las nuevas posibilidades asociadas al *crowdsourcing* y su utilización en el rubro del transporte público. Luego se procede a investigar sobre las plataformas colaborativas y su principal desafío, la baja participación.

2.1 Sistemas inteligentes de transportes (ITS⁷)

Si bien los sistemas inteligentes de transportes fueron utilizados por primera vez como concepto en Estados Unidos en 1986, en donde se discutía los desafíos futuros del transporte de aquel país (Sussman, 2005), no fue hasta el año 2010 que los directivos del Parlamento Europeo y del Consejo definieron los ITS como “aplicaciones avanzadas que sin incorporar inteligencia como tal, pretenden entregar servicios innovativos asociados a diferentes modos de transporte y gestión de tráfico y permitir a varios usuarios a estar mejor informados,..., y a usar más inteligentemente la red de transporte” (EUR-LEX, 2010), reduciéndola a temática asociada transporte rodoviario.

Los ITS tienen seis áreas funcionales: Sistemas de Gestión de Tráfico Avanzado (ATMS), Sistemas de Control Avanzado de Vehículos (AVCS), Operaciones Comerciales de Vehículos (CVO), Sistemas Avanzados de Transporte Rural (ARTS), Sistemas Avanzados de Transporte Público (APTS) y Sistemas Avanzados de Información al Viajero (ATIS) (Sussman, 2005). Las últimas dos están relacionadas al transporte público de la ciudad, en donde el primero busca incrementar la eficiencia del sistema de transporte al aumentar su confiabilidad, y el segundo pretende mejorar la toma de decisiones de los viajeros al desplegar información relacionada a su viaje (Singh & Gupta, 2015).

Los beneficios de los Sistemas Avanzados de Información al Viajero (ATIS) han sido profundamente estudiados. Por ejemplo, el sistema *Bus Tracker* en Chicago demostró que el envío de información de ubicación en tiempo real y tiempo de espera estimado a los usuarios aumentó los viajes en buses en un estudio de tres años (Tang & Thakuriah, 2012), resultados similares a los obtenidos en la aplicación *Bus Time* en Nueva York (Brakewood, Macfarlane, & Watkins, 2015). En cambio en Londres, el despliegue del tiempo de llegada del bus por el sistema *Countdown* provocó tener un impacto positivo en el valor esperado del tiempo de espera en los usuarios de

⁷ “Intelligent Transportation Systems”, en inglés. El resto de los acrónimos en paréntesis del capítulo también están representados en inglés.

buses (Lu, y otros, 2018). Todos estos resultados también fueron obtenidos a partir del estudio de los beneficios de la aplicación *OneBusAway*, donde además se encontró un fuerte incremento en la seguridad y en la satisfacción de los usuarios respecto al transporte público, y beneficios de salud (Ferris, Watkins, & Borning, 2010).

Si bien en los estudios previamente mencionados el flujo de información es desde las autoridades de la gestión del transporte público hacia los usuarios, ha surgido una nueva forma de obtener datos denominada *crowdsourcing*, que es unir piezas de información de muchas personas (Bovornkeeratiroj, Nakorn, & Rojviboonchai, 2015). Debido a la creciente adopción de teléfonos inteligentes por las personas, estos dispositivos podrían ser considerados como herramientas de intercambio de información en *crowdsourcing* (Nandan, Pursche, & Zhe, 2014).

En los sistemas explícitos, que son sistemas independientes donde los usuarios ejecutan una actividad específica, existen cuatro tipos de acciones básicas del *crowdsourcing*: evaluar, compartir, construir artefactos y ejecutar tareas específicas (Misra, Gooze, Watkins, Asad, & Le Dantec, 2014). De las cuales dos se realizan en *Transapp* (evaluar viajes y estado de infraestructura mediante reportes, y compartir información de GPS), por lo que sumada la definición anterior se puede considerar como una aplicación que utiliza el enfoque de *crowdsourcing* en el rubro del transporte.

Actualmente, hay dos aplicaciones que lideran mundialmente el *crowdsourcing* en transporte: *Waze* (transporte privado) y *Moovit* (transporte público). Ambas ayudan a sus usuarios a realizar mejores decisiones de su viaje mediante la entrega de información obtenida de una red de usuarios, por ejemplo: los tiempos de llegada, la congestión, y el hacinamiento y limpieza del bus o tren (sólo en *Moovit*) (Faber & Matthes, 2016).

Debido a la facilidad de la obtención de los datos, el *crowdsourcing* en transporte es útil en países en vías de desarrollo donde no hay información de la oferta de buses (ubicación de éstos, tiempos de llegada) y/o de la demanda (cargas de pasajero, tiempos de espera) del transporte público (Nandan, Pursche, & Zhe, 2014). Por otro lado, también permiten mejorar la calidad del servicio entregado por los proveedores del servicio de transporte público y la efectividad de los sistemas de transportes (Filippi, Fusco, & Nanni, 2013).

Para las aplicaciones *crowdsourcing* es crucial contar con una comunidad *online* de usuarios comprometida (Filippi, Fusco, & Nanni, 2013), cuyos principales desafíos de los desarrolladores para sostener la aplicación son: (1) atraer usuarios e identificar sus necesidades, (2) brindar y solicitar ayuda (entre usuarios), (3) encontrar un balance en la efectividad de los incentivos, (4) controlar efectivamente la presión subjetiva del tiempo fuera del ambiente de la aplicación, y (5) comprender la comunidad (Dickinson, y otros, 2015).

Dentro de las motivaciones de las personas para participar de actividades de *crowdsourcing*, Vianna, Peinado & Graeml (2017) indican 9 factores: premios económicos, reconocimiento y/o

gloria, oportunidades profesionales, conocimiento y desarrollo de habilidades, socializar y/o formar relaciones, placer y/o diversión, interés propio, altruismo y/o deseo de compartir, y amor. En *Transapp*, dado la fisonomía de las herramientas de la aplicación, de éstos factores se considera posible los relacionados a reconocimiento y/o gloria (debido a la *gamificación* en la aplicación) y el altruismo.

2.2 Participación en plataformas online

Tal como se indicó en el capítulo introductorio, *Transapp* también es considerada una plataforma *online* colaborativa y de *crowdsourcing*. En éstas la participación, en general, clasifica dos tipos de usuarios: quienes *postean* regularmente son llamados “*posters*” (*posteadores* en español) y quienes utilizan la información que éstos aportan, sin contribuir a la comunidad *online*, se llaman “*lurkers*” (merodeadores en español) (Baym, 1993) (esto a pesar de no existir un consenso claro en métricas para clasificación (Malinen, 2015)).

Tal como indican Ling *et al.* (2006), en diversas aplicaciones y sitios web se observan porcentajes de participación bastante bajos sobre la proporción de usuarios *posters*, incluso cercanas a cero. Por ejemplo, la aplicación comunitaria de tráfico en automóviles, *Waze*, en 2013 contaba con 50 millones de usuarios mundiales de los cuales sólo el 0,01% reportaba información en la aplicación sobre desvíos u otra información relacionada a tráfico (Weitzenkorn, 2013). De igual manera, un estudio realizado por uno de los fundadores de *Wikipedia* encontró que el 50% de todas las ediciones de contenido de aquella plataforma fueron realizadas por sólo el 0,7% de los usuarios (Swartz, 2006).

Este fenómeno fue condensado por Nielsen (2006) quien elaboró una regla denominada 90-9-1, que indica que el 90% de los usuarios son *lurkers* (es decir, no contribuyen aportando información), el 9% son activos ocasionalmente y el 1% restante contribuye constantemente tanto así que sus aportes consideran la mayoría de las contribuciones. Dentro de los problemas de esta situación, menciona una baja representación de la muestra en caso de *feedback* de consumidores y/o reseñas, entre otros.

En *Transapp* esta gran diferencia en la proporción de *lurkers* y *posters* también está presente, por ejemplo, en marzo de 2019 sólo el 11,6% de los usuarios reportó alguna vez⁸ y es más crítico considerando que el 1% de los usuarios concentra el 53% de todos los reportes realizados en aquel mes, lo cual es bastante similar a la regla de Nielsen. Luego, se considera importante disminuir la cantidad de *lurkers* y aumentar los *posters* para sostener la viabilidad y el éxito de la plataforma (Burke, Marlow, & Lento, 2009) (Malinen, 2015) en *Transapp*.

⁸ Métrica que se usará para ser considerado “reportador/*poster*”.

Nonnencke, Andrews, Preece & Voutour (2004) estudiaron más de 350 plataformas *online* para comprender por qué los *lurkers* no contribuyen, e indican como principales razones: “sólo leer/buscar es suficiente” (54%), “aún estoy aprendiendo de la comunidad” (30%), “timidez de postear” (28%), “no tengo nada que ofrecer” (23%), entre otros. Estos son válidos en *Transapp*, sumados a otros menos relevantes, como: “Otros responden lo que yo diría” (18%), “No me da valor” (11%), “No sé cómo hacerlo” (8%).

Por otro lado, para incentivar que los lurkers comiencen a participar, Sun, Pei-Luen Rau & Ma (2014) mencionan las siguientes estrategias: entregar estímulos externos (es decir, premios tangibles, como regalos o dinero, entre otros), mejorar la usabilidad de la plataforma, animar los usuarios mediante información motivante, y entregar guía a nuevos integrantes para que comprendan la cultura de la plataforma.

2.3 Motivaciones para participar en plataformas online

A continuación se presentan algunos estudios para comprender más de las motivaciones de las personas para participar en las plataformas *online*, a pesar de que éstas podrían diferir de la temática principal de transporte público de la aplicación en la cuál se estudia en el presente trabajo de título.

Serra-Cantallops & Salvi (2014) hicieron una búsqueda de la investigación en *eWOM*⁹ (*electronic word-of-mouth*) y analizaron qué factores motivaban a que los consumidores escribiesen *eWOM* y qué impactos tenía en la industria hotelera. En particular, indican que los principales factores para que las personas escriban reviews de hoteles son: la calidad de servicio, satisfacción e insatisfacción del cliente, el sentimiento de pertenencia a la comunidad *online*, su identidad social, las expectativas previas a la compra, si presenciaron fallas y sus posteriores recuperaciones, si desean ayudar a la empresa y a los otros viajeros. Respecto a esta última motivación, según Bronner & Hoog (2011), es la razón por la cual el 70% de los usuarios participan en una encuesta de vacaciones holandesa.

Profundizando en la motivación por ayudar, una investigación realizada por Ling *et al.* (2006) estudió el *Collective Effort Model* mediante la realización de experimentos de campo en el sitio web *MovieLens* (que presentaba un problema similar de baja contribución), donde los usuarios podían realizar evaluaciones de películas (con las cuales recomiendan a otros usuarios con similares gustos). Para evaluar si las personas se motivan a participar cuando perciben el valor de sus contribuciones, encontraron que mensajes vía mail que destacan la contribución en beneficio propio reducían significativamente la cantidad de evaluaciones; y que el mensaje destacando ambos

⁹ *WOM* es definida por Litvin, Goldsmith & Pan (2008) como “la comunicación entre consumidores acerca de un producto, servicio o empresa, en el cual las fuentes son independientes de la influencia comercial”. Por consecuencia, también indican que *eWOM* es la discusión por medio de internet entre usuarios.

beneficios (a otros e individual) no era significativo. Por otro lado, no encontraron resultados significativos al incentivar la contribución mediante el recordatorio de motivación intrínseca¹⁰. Todos estos resultados mencionados fueron contraproducente con el marco teórico que proponían los autores, que justificaron en intervenciones débiles. A partir de estos hallazgos, Rashid *et al.* (2006) trabajaron en la misma plataforma, *MovieLens*, pero esta vez modificando la interfaz de manera más efectiva, y así obtuvieron que destacar la utilidad personal disminuía significativamente la probabilidad de realizar una evaluación (un 4,8% menos), también que las personas prefieren evaluar si el beneficio es para un grupo en particular (a diferencia que a todos) (un 2,1% más) y que este efecto es incluso más grande si el grupo es de preferencias similares al individuo (un 7,4% más).

Posterior a estos trabajos, Suhonen, Lampinen, Cheshire & Antin (2010) indican, en base a encuestas en una plataforma *online* de intercambio de objetos y servicios, que los usuarios sí quieren saber la utilidad de sus actividades, incluso si es baja. Lo cual podría generar nuevas ideas para plataformas *online* sobre cómo aumentar la participación destacando la utilidad de los aportes (como por ejemplo a los usuarios).

En *Wikipedia*, cuyo beneficio es un bien público al resto de los usuarios (Park, Sook Oh, & Kang, 2015), Park *et al.* (2015) estudiaron las motivaciones de los individuos para subir contenido mediante encuestas, las cuales analizaron usando un modelo estructural. Encontraron evidencia que la actitud, las normas subjetivas, el control percibido de comportamiento y las normas descriptivas son significativas y están positivamente asociadas a la intención de subir contenido. De igual manera, encontraron que el ego también está relacionado positivamente con la intención de subir contenido a *Wikipedia*, a través de la actitud, la intención y el control percibido de comportamiento. Dentro de las limitaciones del estudio enuncian la baja muestra del estudio y la baja validez externa al ser sólo estudiantes universitarios de un centro educativo.

Asimismo Dixit, Badgaiyan & Khare (2019) realizaron un estudio para investigar las motivaciones de las personas para escribir reseñas *online* de restaurantes utilizando una metodología similar al estudio previo. Sus resultados consisten en que: el ego, el sentimiento de “vengarse” del local, las normas subjetivas y el control percibido de comportamiento son predictores significativos de la probabilidad de realizar una reseña, a diferencia de la intención de ayudar al restaurante y la actitud, que resultaron ser no significativas. Los autores proponen estudios futuros que utilicen nuevos sectores comerciales y métodos de investigación mediante el desarrollo de experimentos que permitan testear cambios en las variables dependientes e independientes.

¹⁰ Este incentivo mencionaba la diversión asociada a la acción de realizar una evaluación de una película. Para una revisión más precisa de la definición de motivación intrínseca, esta se discute en el subcapítulo 5.6.3 Conclusiones teóricas del experimento realizado.

Capítulo 3

Análisis de datos

En este capítulo se analiza las bases de datos mensuales de *Transapp*, las cuales indican las acciones realizadas por los usuarios en la aplicación, con espacio temporal de Agosto de 2018 a Septiembre de 2019. A continuación, en la Tabla 1, se presenta los principales atributos de cada una.

Tabla 1: Descripción de las bases de datos de *Transapp* a analizar.

| Nombre de la base de datos | Contenido almacenado |
|-----------------------------------|---|
| Consultas de paradero de usuarios | Contiene las consultas realizadas por los usuarios en los paraderos de la ciudad para saber el tiempo de llegada de los siguientes buses. |
| Reportes de usuarios | Indica los reportes de los usuarios sobre el sistema de transporte, ya sean sobre buses o paraderos. |
| Viajes de usuarios | Almacena los viajes realizados por los usuarios en la aplicación, y si finalmente evaluaron (o no) el viaje. |

En este capítulo se utilizan frecuentemente ciertos términos asociados a los usuarios, por lo que en el siguiente glosario (ver Tabla 2) se procede a definirlos para facilitar la lectura del documento.

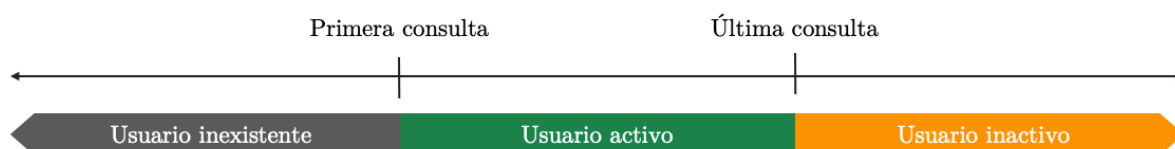
Tabla 2: Glosario de términos frecuentes asociados a usuarios.

| Término | Significado |
|--------------------|--|
| Usuario consultor | Usuario que consulta tiempos de llegada de buses en los paraderos. |
| Usuario reportador | Usuario que reporta sobre alguna infraestructura. |
| Usuario viajador | Usuario que viaja en un recorrido de bus. |
| Usuario evaluador | Usuario que viaja y evalúa un recorrido de bus. |

3.1. Usuarios de la aplicación según consultas de paraderos

En este capítulo se considera un usuario como activo en un determinado momento, si en aquel instante está dentro del intervalo entre la primera y última consulta registrada acerca del tiempo de llegada de buses en un paradero. Se utiliza este intervalo ya que se presume que el usuario podría potencialmente haber utilizado la aplicación, en cambio después de la última consulta se presume que el usuario deja de utilizar la aplicación y se convierte en inactivo. En la Ilustración 3 se explica visualmente la clasificación.

Ilustración 3: Explicación de usuario activo e inactivo en la aplicación, en base a la consulta de paraderos.



Fuente: elaboración propia

En la Tabla 3, se muestra las variaciones de los usuarios activos mensual. Los usuarios nuevos y “de baja” (que se convierten en inactivos) se miden respecto al mes previo, así como el delta de usuarios y la variación porcentual de usuarios activos.

Tabla 3: Cantidad mensual de usuarios activos, nuevos, “de baja” y la variación mensual.

| Año | Mes | Usuarios activos | Usuarios nuevos | Usuarios “de baja” | Delta de usuarios | Variación porcentual |
|------|------------|------------------|-----------------|--------------------|-------------------|----------------------|
| 2018 | Agosto | 13.504 | - | - | - | - |
| | Septiembre | 17.939 | 8.214 | 3.779 | 4.435 | 32,84% |
| | Octubre | 21.741 | 9.149 | 5.347 | 3.802 | 21,19% |
| | Noviembre | 24.924 | 9.882 | 6.699 | 3.183 | 14,64% |
| | Diciembre | 28.329 | 10.267 | 6.862 | 3.405 | 13,66% |
| 2019 | Enero | 31.109 | 10.678 | 7.898 | 2.780 | 9,81% |
| | Febrero | 30.971 | 7.992 | 8.130 | -138 | -0,44% |
| | Marzo | 39.087 | 14.568 | 6.452 | 8.116 | 26,21% |
| | Abril | 41.120 | 12.579 | 10.546 | 2.033 | 5,20% |
| | Mayo | 40.975 | 10.457 | 10.602 | -145 | -0,35% |
| | Junio | 40.940 | 9.703 | 9.738 | -35 | -0,09% |
| | Julio | 41.345 | 9.183 | 8.778 | 405 | 0,99% |
| | Agosto | 42.921 | 10.719 | 9.143 | 1.576 | 3,81% |
| | Septiembre | 47.320 | 15.417 | 11.018 | 4.399 | 10,25% |

Por otro lado, se tiene que la cantidad promedio de semanas de actividad es de 10,28, y presenta una distribución bastante irregular respecto a la proporción de usuarios, en donde el 37,41% de los usuarios tiene sólo una semana de actividad y sólo el 7,39% tiene más de seis meses. Esta se observa en la Ilustración 4 y en la Tabla 4.

Ilustración 4: Distribución de usuarios según número de semanas de actividad.

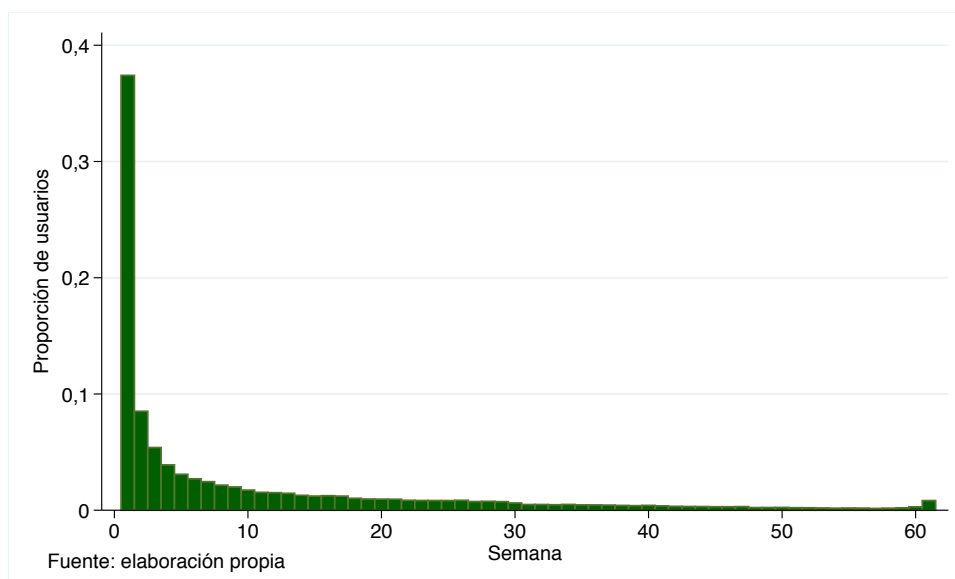


Tabla 4: Cantidad y proporción de usuarios activos en determinadas semanas, considerando toda la muestra.

| Semanas de actividad | Cantidad de usuarios | Proporción de usuarios ¹¹ | Acumulado |
|-----------------------|----------------------|--------------------------------------|-----------|
| Sólo 1 | 56.976 | 37,41% | 37,41% |
| Entre 2 y 4 | 27.114 | 17,80% | 55,21% |
| Entre 5 y 12 | 26.168 | 17,18% | 72,39% |
| Entre 13 y 24 | 19.520 | 12,82% | 85,21% |
| Entre 25 y 36 | 11.276 | 7,40% | 92,61% |
| Más de 37 | 11.258 | 7,39% | 100,00% |
| Total usuarios | 152.312 | 100,00% | - |

Si bien la Tabla 4 entrega una panorámica de la actividad de los usuarios en *Transapp*, para poder concluir acerca de cuántos usuarios siguen utilizando la aplicación después de determinada cantidad de semanas, se procede a eliminar los usuarios que no tuvieron el tiempo mínimo dependiendo del corte utilizado. Por ejemplo, si un usuario empezó a utilizar *Transapp* en la última

¹¹ Este porcentaje es respecto de la cantidad de usuarios totales existentes que han consultado algún paradero durante el intervalo de estudio. No omite a los usuarios que les sería imposible, por ejemplo, llegar a las 36 (u otra cantidad de) semanas. Esto en el caso de que alguien que ingresa en la semana 60, no podría ser un usuario de más de 36 semanas, dado que sólo tiene 2 de actividad.

semana del estudio, lógicamente no superaría las 4 semanas de actividad, por lo que se resta de este análisis. El cual, extendido al resto de las unidades temporales, se adjunta en la Tabla 5.

Tabla 5: Cantidad y proporción de usuarios activos en determinadas semanas, considerando la muestra factible.

| Unidad temporal | Usuarios con actividad | | | |
|-------------------------|------------------------------------|-------------------------------|----------------------------|-------------------------------|
| | Menor o igual a la unidad temporal | Proporción del total factible | Mayor a la unidad temporal | Proporción del total factible |
| 1 semana | 52.230 | 43,5% | 67.895 | 56,5% |
| 4 semanas (1 mes) | 66.672 | 65,1% | 35.673 | 34,9% |
| 12 semanas (3 meses) | 72.438 | 85,4% | 12.393 | 14,6% |
| 24 semanas (6 meses) | 56.851 | 97,4% | 1.515 | 2,6% |
| 36 semanas (9 meses) | 33.869 | 100,0% | 0 | 0,0% |

De la Tabla 5 se tiene que 66.672 usuarios utilizaron activamente la aplicación menos de 1 mes (menor o igual a 4 semanas), equivalente al 65,1%. De igual forma, 35.673 usuarios la utilizaron por más de 1 mes, completando el 34,9% restante. Así se tiene que sólo el 56,5% de los usuarios continúa consultando tiempos de llegada en la aplicación posterior a la primera semana de uso, cifra que disminuye al 14,6% a los 3 meses. Esto indica que, si bien la cantidad de usuarios de la aplicación crece de manera relativamente regular, existe un constante recambio de usuarios, ya que sólo el 2,6% de los usuarios totales ha estado activo consultando más de 6 meses y ninguno más de 9.

3.2. Análisis de consultas de paraderos

La consulta de paraderos es la actividad más realizada por los usuarios en *Transapp*, y dado que su base de datos almacena cada consulta realizada, es la más voluminosa. Con esta información primero se realiza un análisis descriptivo general, y uno local enfocado en características horarias.

3.2.1 Análisis descriptivo

3.2.1.1 Estadísticas básicas de consultas de paraderos

En la Tabla 6, se incluye la cantidad de consultas a paraderos realizada, la cantidad de consultas a paraderos únicas, la cantidad de usuarios activos, la cantidad de usuarios consultores, la proporción que éstos simbolizan de los usuarios potencialmente activos y la cantidad promedio de consultas por usuario que consulta.

Se realiza dos diferencias: (1) entre consultas realizadas y consultas únicas realizadas, ya que en Abril de 2019 se realizó un cambio a la forma de almacenamiento al número de consultas involuntarias (aquellas que luego de consultar, se actualizaba automáticamente cada 30 segundos). La métrica utilizada fue declarar como consulta única a aquellas consultas del mismo paradero en un intervalo de 30 minutos. (2) Entre usuarios activos y consultores, un usuario puede estar activo en un período pero no necesariamente haber consultado.

Tabla 6: Estadísticas descriptivas de las consultas de paraderos mensuales.

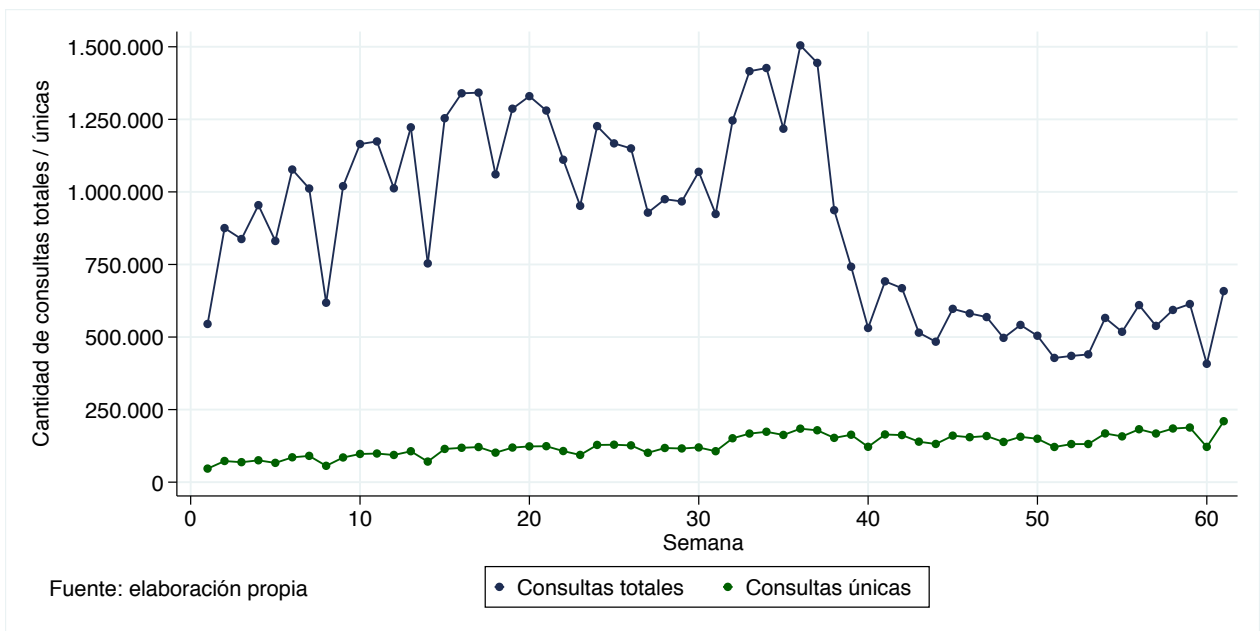
| Año | Mes | Consultas realizadas | Consultas únicas | Usuarios activos | Usuarios consultores | Proporción usuarios consultores | Promedio consultas únicas por usuario consultor |
|------------|------------|----------------------|------------------|------------------|----------------------|---------------------------------|---|
| 2018 | Agosto | 3.828.720 | 312.975 | 13.504 | 13.504 | - | 23,18 |
| | Septiembre | 3.940.819 | 333.375 | 17.939 | 17.303 | 96,45% | 18,58 |
| | Octubre | 4.933.152 | 429.087 | 21.741 | 20.733 | 95,36% | 19,74 |
| | Noviembre | 5.123.579 | 466.926 | 24.924 | 23.459 | 94,12% | 18,73 |
| | Diciembre | 5.273.057 | 498.296 | 28.329 | 26.375 | 93,10% | 17,59 |
| 2019 | Enero | 5.026.956 | 534.373 | 31.109 | 28.478 | 91,54% | 17,18 |
| | Febrero | 3.867.438 | 453.010 | 30.971 | 26.937 | 86,97% | 14,63 |
| | Marzo | 5.769.980 | 705.023 | 39.087 | 36.275 | 92,81% | 18,04 |
| | Abril | 4.738.263 | 703.963 | 41.120 | 37.938 | 92,26% | 17,12 |
| | Mayo | 2.673.967 | 663.068 | 40.975 | 37.474 | 91,46% | 16,18 |
| | Junio | 2.349.894 | 641.884 | 40.940 | 37.181 | 90,82% | 15,68 |
| | Julio | 2.073.002 | 605.798 | 41.345 | 37.708 | 91,20% | 14,65 |
| | Agosto | 2.454.104 | 742.151 | 42.921 | 40.637 | 94,68% | 17,29 |
| Septiembre | 2.326.049 | 719.545 | 47.320 | 47.320 | - | 15,21 | |

En la tabla se observa cómo la cantidad de consultas realizadas en el comienzo crece relativamente estable en niveles de 4 y 5 millones mensuales, no obstante en el mes de Abril hay una fuerte disminución que finaliza en una estabilización en niveles entre 2 y 2,5 millones, los cuales coinciden con el período en el cual el equipo de *Transapp* modificó el almacenamiento de la data.

Por otro lado, se observa que la cantidad de consultas únicas también crece, de niveles inferiores a 500 mil consultas en 2018 a por sobre de 600 mil en 2019. Esto también en línea con la cantidad de usuarios activos y usuarios consultores que también aumentan en el período en estudio. Por otro lado, se tiene que la proporción de usuarios activos que consultan se mantiene estable en el rango de 90-95%.

Para realizar un análisis más preciso, se utiliza datos semanales para así igualar la cantidad de días, ya sea laborales y de fin de semana. La tabla está en Anexos, en el Apéndice A: Estadísticas de consultas de paradero por semana, de la cual se obtiene los siguientes gráficos (Ilustración 5 e Ilustración 6).

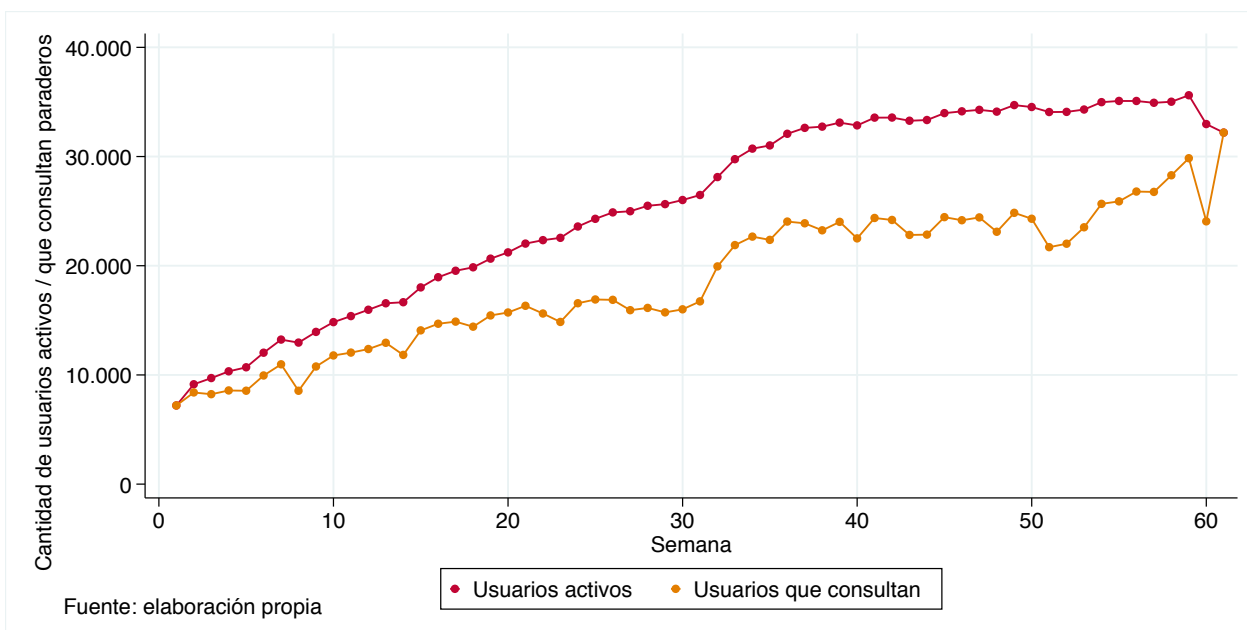
Ilustración 5: Distribución semanal de las consultas realizadas y las consultas únicas.



Se confirma lo mencionado, ya que si bien la cantidad de consultas totales en el período de la semana 37-39 (durante Abril de 2019) disminuye, la cantidad de consultas únicas sigue un patrón más estable, esto debido al filtro realizado (conversión de consultas realizadas a únicas).

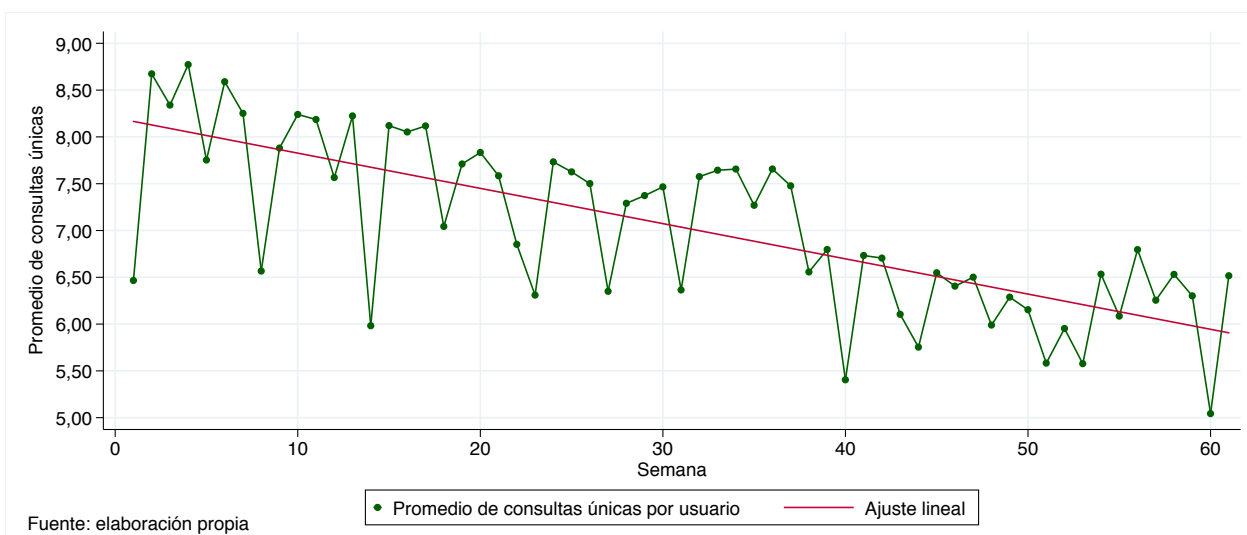
En la Ilustración 6 se observa el crecimiento sostenido de la cantidad de usuarios en la aplicación y cómo la cantidad de usuarios consultores sigue la misma dirección. En la Tabla 6, se muestra niveles de proporción de usuarios activos mensuales cercanos a 90%, que promedian 93,63%, el cual es mayor al 73,89% semanal. Esto es esperable ya que disminuye el espacio temporal para que un usuario activo consulte un paradero.

Ilustración 6: Número de usuarios activos y usuarios consultores, semanalmente.



En los dos gráficos previos, Ilustración 5 e Ilustración 6, se muestra un crecimiento en la cantidad de consultas de paraderos y los usuarios consultores, luego es interesante ver cómo ha variado la cantidad promedio de consultas por usuario consultor semanalmente, la cual se grafica a continuación en la Ilustración 7.

Ilustración 7: Promedio de consultas por usuario consultor, semanalmente.



En esta se observa una tendencia negativa, la cual se visibiliza por el ajuste lineal, lo que indica que cada vez hay menos consultas únicas por usuario semanalmente. Particularmente, en las 10 primeras semanas los valores fluctúan entre 6,5 y 9 consultas promedio; mayores a los observados en las últimas 10, cuyos valores fluctúan entre 5 y 7.

3.2.1.2 Estadísticas de usuarios consultores

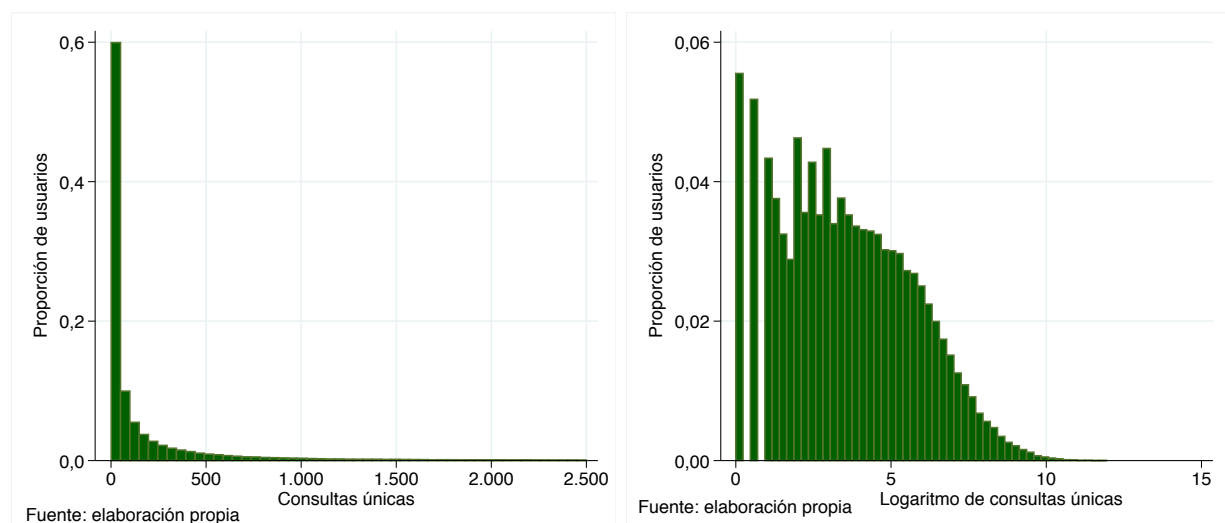
Al considerar la cantidad total de consultas únicas realizadas en el período se presenta la siguiente distribución, en donde: un tercio de los usuarios acumula menos de 10 consultas únicas, el 80% de los usuarios acumula menos de 250, y sólo el 3% acumula una cantidad superior a 2.500.

Tabla 7: Cantidad y proporción de usuarios que realizan consultas únicas en todo el período.

| Cantidad de consultas únicas | Cantidad de usuarios | Proporción de usuarios | Acumulado |
|------------------------------|----------------------|------------------------|-----------|
| Sólo 1 | 8.456 | 5,55% | 5,55% |
| Entre 2 y 10 | 42.028 | 27,59% | 33,15% |
| Entre 11 y 50 | 38.740 | 25,43% | 58,58% |
| Entre 51 y 250 | 32.453 | 21,31% | 79,89% |
| Entre 251 y 1.000 | 19.964 | 13,11% | 92,99% |
| Entre 1.001 y 2.500 | 6.506 | 4,27% | 97,27% |
| Más de 2.501 | 4.165 | 2,73% | 100,00% |
| Total usuarios | 152.312 | 100,00% | - |

En el gráfico de la izquierda de la Ilustración 8, se ilustra la desigual proporción de usuarios según la cantidad de consultas únicas, concentrada sólo en aquellos con menos de 2.500 consultas (el 97% de los usuarios). Debido a la existencia de usuarios que consultan sobre 50.000 veces, para facilitar la visualización de la distribución se presenta el gráfico de la derecha que considera el logaritmo de consultas únicas realizadas, y cómo esta decae progresivamente en los valores mayores a 7,5 ($\exp(7,5)=1.808$ consultas), que son los usuarios no considerados en el gráfico de la izquierda.

Ilustración 8: Histograma de usuarios según consultas únicas (izquierda) y logaritmo de consultas únicas (derecha).



3.2.2 Análisis local

En esta sección se busca ver cómo varían las consultas de paraderos según el intervalo horario y día de la semana, para así identificar en cuáles momentos hay más personas utilizando la aplicación.

3.2.2.1 Análisis local según horario

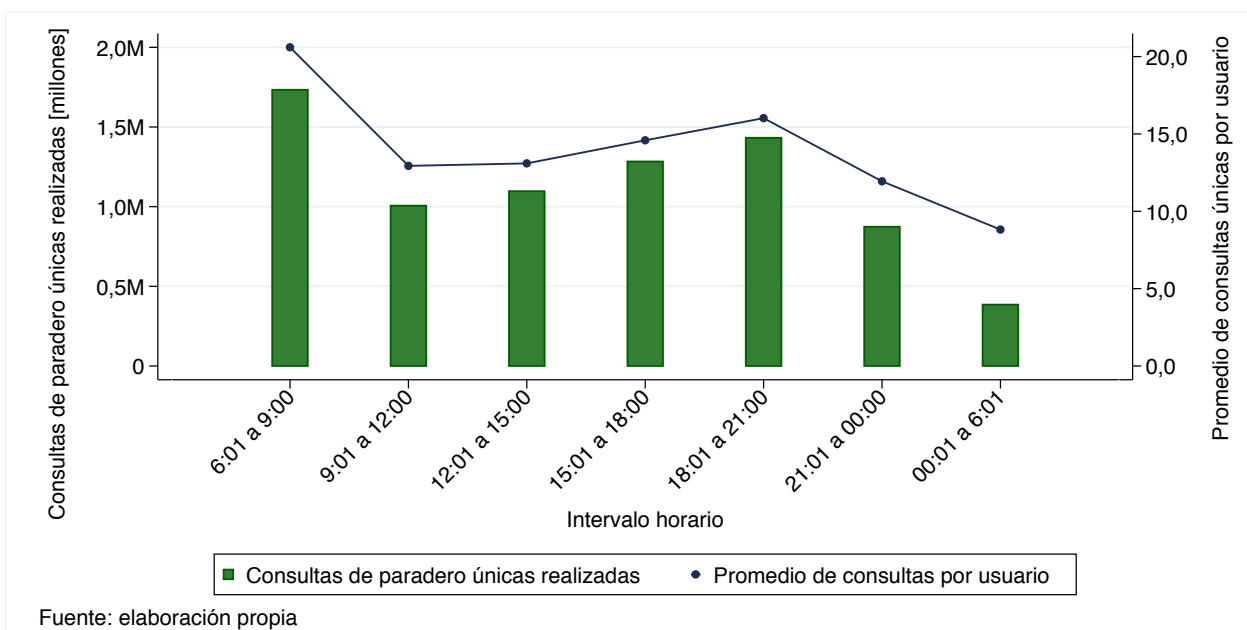
En la Tabla 8, se presenta la cantidad de consultas de paraderos totales en el período según intervalos de 3 horas diarias iniciando a las 6 am, a excepción del período nocturno-madrugada de 00:00 a 6:00 hrs. Estos períodos permiten aislar la hora punta mañana y tarde, la hora valle previo el medio día y posterior a las 15 hrs, y la hora semi-*peak* de almuerzo, que son períodos con gran diferencias entre ellos en término de viajes.

Tabla 8: Cantidad de consultas únicas de paraderos, usuarios consultores y promedio de consultas únicas por usuario según intervalo horario del día.

| Intervalo horario | Consultas únicas realizadas | Porcentaje | Acumulado | Usuarios consultores | Promedio de consultas por usuario |
|-------------------|-----------------------------|----------------|-----------|----------------------|-----------------------------------|
| 6:01 a 9:00 | 1.733.210 | 22,19% | 22,19% | 84.147 | 20,60 |
| 9:01 a 12:00 | 1.006.085 | 12,88% | 35,08% | 77.771 | 12,94 |
| 12:01 a 15:00 | 1.096.839 | 14,04% | 49,12% | 83.754 | 13,10 |
| 15:01 a 18:00 | 1.283.177 | 16,43% | 65,55% | 87.966 | 14,59 |
| 18:01 a 21:00 | 1.431.953 | 18,34% | 83,89% | 89.391 | 16,02 |
| 21:01 a 00:00 | 873.511 | 11,19% | 95,07% | 73.169 | 11,94 |
| 00:01 a 6:00 | 384.699 | 4,93% | 100,00% | 43.636 | 8,82 |
| Total | 7.809.474 | 100,00% | - | - | - |

La Ilustración 9 refleja la información relacionada a la cantidad de consultas únicas y el promedio de consultas por usuario. Si bien se observan dos *peaks* en ambas variables en las respectivas horas punta de la mañana y tarde, la primera es la máxima.

Ilustración 9: Consultas únicas por paradero y promedio de consultas únicas por usuario, según hora del día.



Esta información es útil ya que permite identificar los intervalos horarios en donde hay más personas utilizando efectivamente la aplicación. Por ejemplo, si se desea mandar una notificación masiva, es más probable que más usuarios la vean en el instante (o próximo al envío) si es enviada en la mañana a diferencia de horarios posteriores a la hora punta de la tarde.

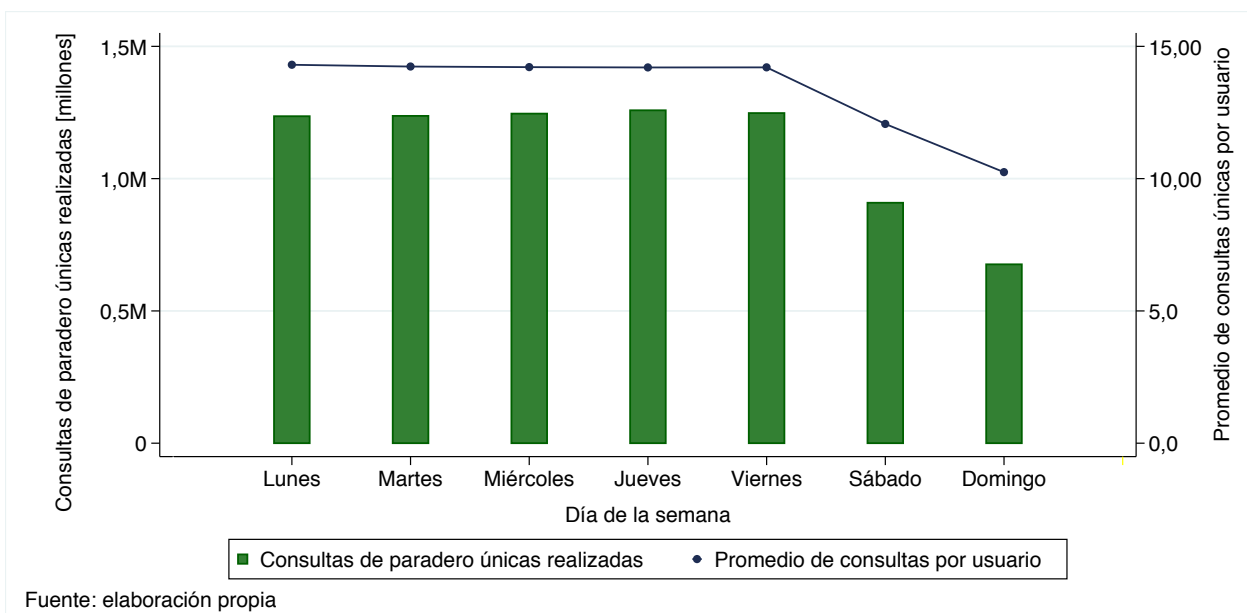
3.2.2.2 Análisis local según día de la semana

En la Tabla 9, se presenta la cantidad de consultas únicas totales (en todo el período) según el día de la semana, así como también la cantidad de usuarios consultores y el promedio de consultas por usuario. El posterior gráfico de la Ilustración 10 utiliza la información de esta tabla.

Tabla 9: Consultas únicas por paradero y promedio de consultas únicas por usuario, según día de la semana.

| Intervalo horario | Consultas únicas realizadas | Porcentaje | Acumulado | Usuarios consultores | Promedio de consultas por usuario |
|-------------------|-----------------------------|----------------|-----------|----------------------|-----------------------------------|
| Lunes | 1.235.960 | 15,83% | 15,83% | 86.412 | 14,30 |
| Martes | 1.237.068 | 15,84% | 31,67% | 86.889 | 14,24 |
| Miércoles | 1.245.624 | 15,95% | 47,62% | 87.627 | 14,22 |
| Jueves | 1.258.271 | 16,11% | 63,73% | 88.596 | 14,20 |
| Viernes | 1.247.740 | 15,98% | 79,71% | 87.830 | 14,21 |
| Sábado | 908.752 | 11,64% | 91,34% | 75.311 | 12,07 |
| Domingo | 676.059 | 8,66% | 100,00% | 65.989 | 10,25 |
| Total | 7.809.474 | 100,00% | - | - | - |

Ilustración 10: Consultas únicas de paradero y promedio de consultas únicas por usuario, según día de la semana.



Se desprende que la cantidad total de consultas únicas es casi idéntica en los días laborales, sin embargo, desciende progresivamente en los sábados y domingos. Notar que si bien la cantidad de consultas únicas de los domingos es la mitad de un día laboral (reducción de 50%), la cantidad promedio de consultas por usuario sólo se reduce un 28% (de aproximadamente 14,25% a 10,25%).

3.3. Análisis de reportes

Primero, hay que tener en cuenta que en categoría de reportes existe una fuerte categorización de usuarios en términos de que hay gran cantidad que reporta poco, y una baja cantidad que reporta mucho (esta situación se aborda más en profundidad en el subcapítulo: Problema de participación). Dentro de esta última categoría existe un tipo de usuario particular que reporta de manera *excesiva*, incluso a veces reportando de infraestructura faltando a la objetividad. Luego, este tipo de usuario en vez de ayudar, genera un problema pues declina reportes verídicos o reporta aspectos falsos, con un objetivo, que se presume, es reportar más para así aumentar su nivel en la *gamificación*.

Para identificar estos usuarios que sobre-reportan, se filtra a quienes reportaron sobre una misma infraestructura (recorrido/patente de bus o paradero) más de 20 veces en un mismo día. Luego, todos los reportes de éstos usuarios son eliminados. La razón de esta cantidad radica en que la máxima cantidad de reportes posibles para un paradero/bus son 13. La diferencia de 7 reportes es por si el usuario acepta/niega un reporte previo, doblando así los reportes realizados sobre el mismo atributo; o si reportaba sobre el mismo paradero en diferentes períodos del día.

Utilizando este filtro se identifica a 142 usuarios, que en conjunto suman 250.123 reportes¹², promediando 1.761 cada uno y 104 diarios¹³. Se destaca que estos individuos, equivalentes al 0,34% de los usuarios reportadores de la base original, reportaron el 35,87% de los reportes.

Tabla 10: Filtro de la base de datos de reportes.

| Datos | Número de reportes | Porcentaje | Usuarios | Porcentaje |
|--------------------|--------------------|---------------|---------------|---------------|
| Original | 697.206 | 100,00% | 41.708 | 100,00% |
| Filtrados | 250.123 | 35,87% | 142 | 0,34% |
| “Correctos” | 447.083 | 64,13% | 41.566 | 99,65% |

En la Tabla 10 se le asigna el nombre (reportes) “Correctos” a la base de datos con la cual se realiza el análisis de éste capítulo. Se deja en claro que no es posible determinar objetivamente cuál reporte registrado es correcto o incorrecto, o cuántos reportes diarios es *creíble* que un individuo realice. Se espera sí que este filtro realizado permita haber disminuido los usuarios que reportan negativamente en la aplicación, y así acercarse a los usuarios que utilizan la herramienta de manera consciente y objetiva.

¹² Notar que un solo usuario realizó más de 60.000 reportes.

¹³ Esta cifra es el promedio general del promedio diario de cada uno de los usuarios.

3.3.1 Análisis descriptivo

3.3.1.1 Estadísticas básicas de los reportes realizados

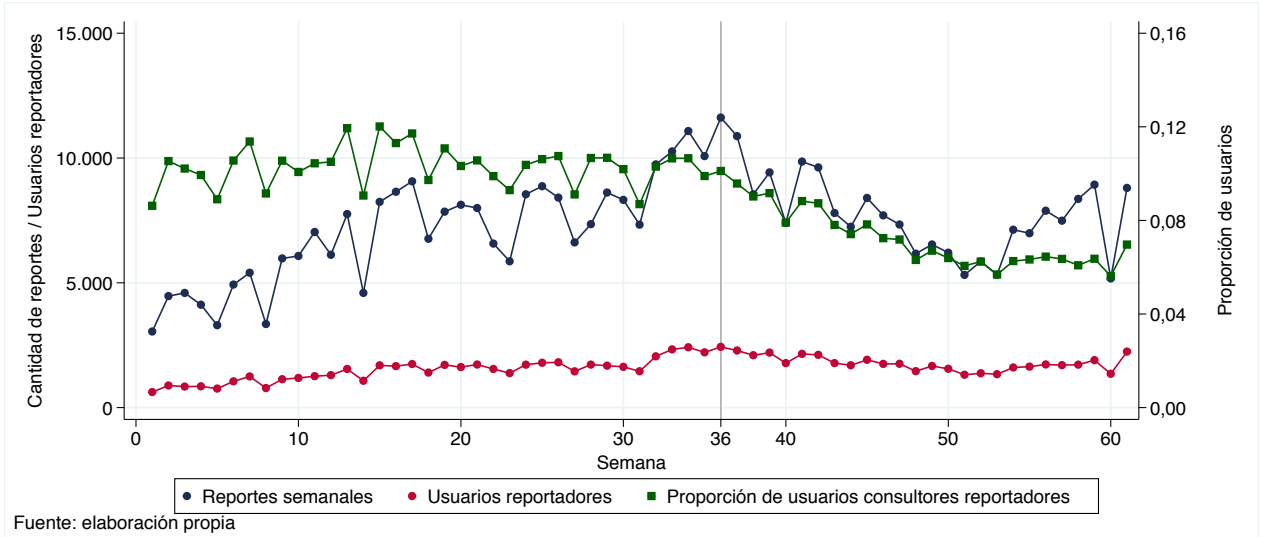
En la Tabla 11 se adjunta la cantidad de reportadores y reportes realizados, y el promedio de reportes por usuario. También se adjunta la cantidad de usuarios activos, obtenida del capítulo anterior, con lo cual se obtiene una proporción de cuántos usuarios activos son reportadores.

Tabla 11: Estadísticas descriptivas de los reportes y reportadores, mensualmente.

| Año | Mes | Usuarios consultores | Usuarios reportadores | Reportes realizados | Promedio de reportes por usuario | Proporción de usuarios reportadores |
|------|------------|----------------------|-----------------------|---------------------|----------------------------------|-------------------------------------|
| 2018 | Agosto | 13.504 | 2.535 | 18.775 | 7,41 | 18,77% |
| | Septiembre | 17.303 | 3.270 | 20.402 | 6,24 | 18,90% |
| | Octubre | 20.733 | 3.960 | 29.236 | 7,38 | 19,10% |
| | Noviembre | 23.459 | 4.687 | 33.736 | 7,20 | 19,98% |
| | Diciembre | 26.375 | 4.868 | 31.870 | 6,55 | 18,46% |
| 2019 | Enero | 28.478 | 5.325 | 35.498 | 6,67 | 18,70% |
| | Febrero | 26.937 | 4.578 | 31.410 | 6,86 | 17,00% |
| | Marzo | 36.275 | 6.772 | 44.179 | 6,52 | 18,67% |
| | Abril | 37.938 | 6.601 | 42.009 | 6,36 | 17,40% |
| | Mayo | 37.474 | 6.043 | 38.984 | 6,45 | 16,13% |
| | Junio | 37.181 | 5.181 | 30.993 | 5,98 | 13,93% |
| | Julio | 37.708 | 4.650 | 25.820 | 5,55 | 12,33% |
| | Agosto | 40.637 | 5.298 | 32.232 | 6,08 | 13,04% |
| | Septiembre | 47.320 | 5.601 | 31.939 | 5,70 | 11,84% |

A continuación se grafica estas variables (ver Ilustración 11 e Ilustración 12), no obstante, en unidad temporal de semanas. La versión semanal de la tabla previa está en Anexos, Apéndice B: Estadísticas de reportes por semana.

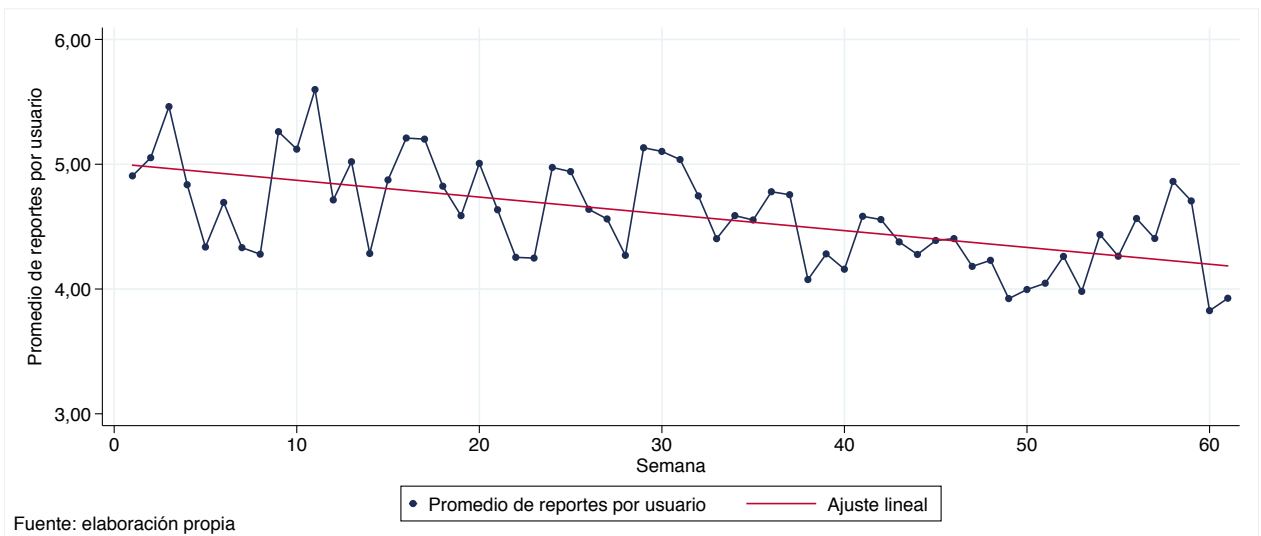
Ilustración 11: Cantidad de reportes semanales, usuarios reportadores y proporción de reportadores, semanalmente.



Se observa en la Ilustración 11 cómo la cantidad de reportes mantiene una tendencia positiva hasta la semana 36, donde alcanza su máximo de reportes, para luego descender a niveles en torno a los 7.500 reportes semanales. Por el otro lado, la cantidad de reportadores sigue un patrón similar, la cantidad es creciente hasta su peak en la semana 36 (línea gris en el gráfico) con 2.431 usuarios, para luego disminuir y mantenerse estable en cercanías de 1.500 usuarios. Se presume que este aumento en la cantidad de reportes se debe al incremento en la cantidad de usuarios activos en la aplicación (ver Ilustración 6), potencialmente causado a publicidad instalada en paraderos.

Al considerar esta información, es interesante ahondar en la cantidad promedio de reportes por usuario reportador, que varía entre 4 y 6 reportes de manera general, y que se grafica en la Ilustración 12, donde se añade un ajuste lineal para identificar una posible tendencia negativa, la cual es visible a pesar de la varianza de los promedios semanales.

Ilustración 12: Promedio de reportes por usuario, semanalmente.



3.3.1.2 Estadísticas de usuarios reportadores

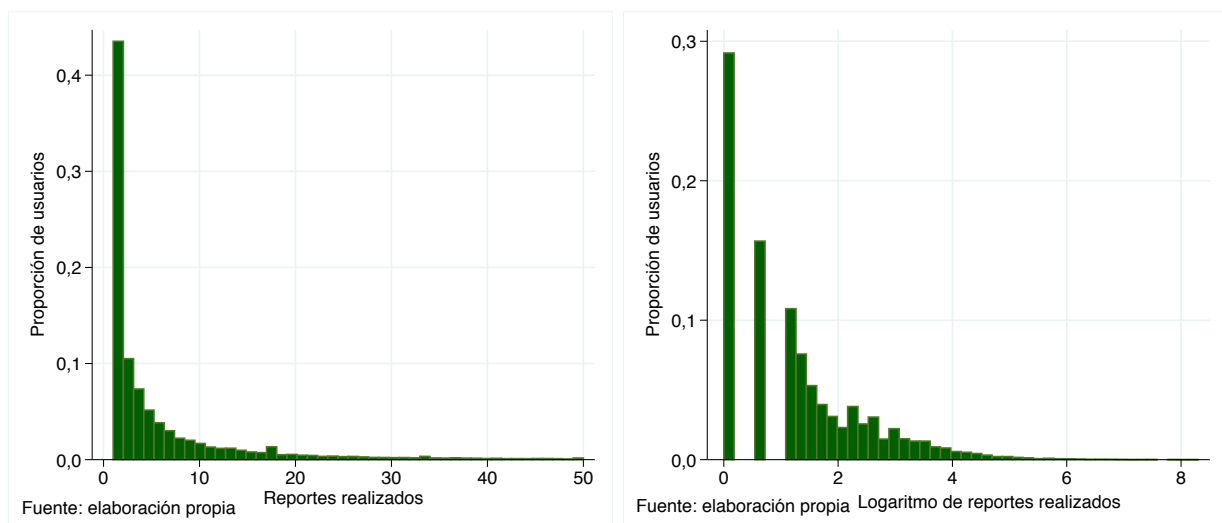
Al considerar la cantidad total de reportes realizados en el período se presenta la siguiente distribución (ver Tabla 12), en donde aproximadamente el 82% de los usuarios tiene menos de 4 reportes realizados. Y donde sólo cerca del 4% reporta más de 26 veces.

Tabla 12: Cantidad y proporción de usuarios que realizan viajes en determinadas semanas.

| Cantidad de reportes | Cantidad de usuarios | Proporción de usuarios | Acumulado |
|-----------------------|----------------------|------------------------|-----------|
| Sólo 1 | 12.121 | 29,16% | 29,16% |
| Entre 2 y 4 | 14.161 | 34,07% | 63,23% |
| Entre 5 y 10 | 7.678 | 18,47% | 81,70% |
| Entre 11 y 25 | 4.494 | 10,81% | 92,51% |
| Entre 26 y 50 | 1.736 | 4,18% | 96,69% |
| Más de 51 | 1.376 | 3,31% | 100,00% |
| Total usuarios | 20.067 | 100,00% | - |

En el gráfico de la izquierda de la Ilustración 13 se adjunta la proporción de usuarios que tienen menos de 50 reportes (el 97% de los usuarios). Debido a la existencia de usuarios con gran cantidad de reportes (por sobre 3.000), para facilitar la visualización se grafica el logaritmo de los reportes realizados en la ilustración de la derecha. En esta se observa cómo en aquellos valores superiores a 4 ($\exp(4)=54,5$ consultas), valores no incluidos en el gráfico de la izquierda, disminuye progresivamente la cantidad de usuarios.

Ilustración 13: Proporción de usuarios según reportes realizados (izquierda) y logaritmo de reportes realizados (derecha).



3.3.2 Análisis local

En esta sección se procede a explicar cómo varía los reportes según atributos propios de éstos: si son acerca de infraestructura de paraderos o buses, si son positivos o negativos; o cómo varían dependiendo del día de la semana y el intervalo horario del día.

3.3.2.1 Reportes de paraderos o buses

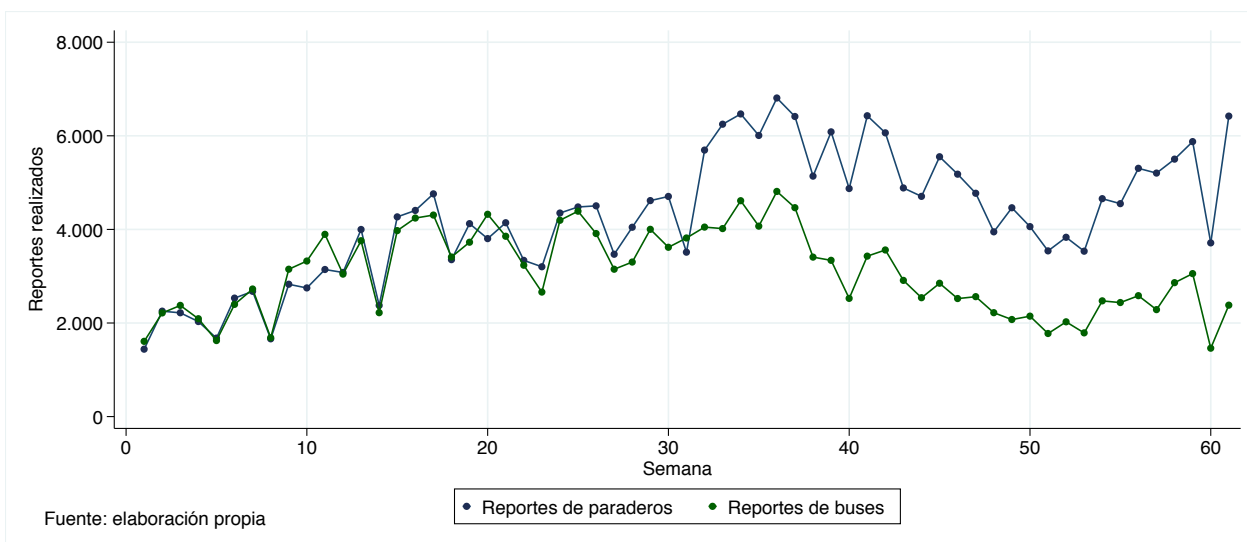
Se procede a realizar un análisis descriptivo de los reportes, especificando sobre qué tipo de infraestructura corresponden. En la Tabla 13 se adjunta la cantidad de reportes mensuales realizados y cuántos son acerca de un paradero y bus, además de la proporción respectiva.

Tabla 13: Cantidad de reportes mensuales, acerca de paradero y bus, mensualmente.

| Año | Mes | Reportes realizados | Reportes de paraderos | Proporción de paraderos | Reportes de buses | Proporción de buses |
|------|------------|---------------------|-----------------------|-------------------------|-------------------|---------------------|
| 2018 | Agosto | 18.775 | 9.188 | 48,94% | 9.587 | 51,06% |
| | Septiembre | 20.402 | 10.123 | 49,62% | 10.279 | 50,38% |
| | Octubre | 29.236 | 14.140 | 48,37% | 15.096 | 51,63% |
| | Noviembre | 33.736 | 17.248 | 51,13% | 16.488 | 48,87% |
| | Diciembre | 31.870 | 16.146 | 50,66% | 15.724 | 49,34% |
| 2019 | Enero | 35.498 | 18.454 | 51,99% | 17.044 | 48,01% |
| | Febrero | 31.410 | 16.903 | 53,81% | 14.507 | 46,19% |
| | Marzo | 44.179 | 25.941 | 58,72% | 18.238 | 41,28% |
| | Abril | 42.009 | 25.371 | 60,39% | 16.638 | 39,61% |
| | Mayo | 38.984 | 25.102 | 64,39% | 13.882 | 35,61% |
| | Junio | 30.993 | 20.377 | 65,75% | 10.616 | 34,25% |
| | Julio | 25.820 | 17.193 | 66,59% | 8.627 | 33,41% |
| | Agosto | 32.232 | 21.456 | 66,57% | 10.776 | 33,43% |
| | Septiembre | 31.939 | 22.002 | 68,89% | 9.937 | 31,11% |

A continuación se grafica esta información (ver Ilustración 14 e Ilustración 15) pero utilizando un intervalo de datos semanal, cuya tabla (similar a la previa) está en Anexos, Apéndice C: Estadísticas de reportes acerca de buses o paradero, por semana.

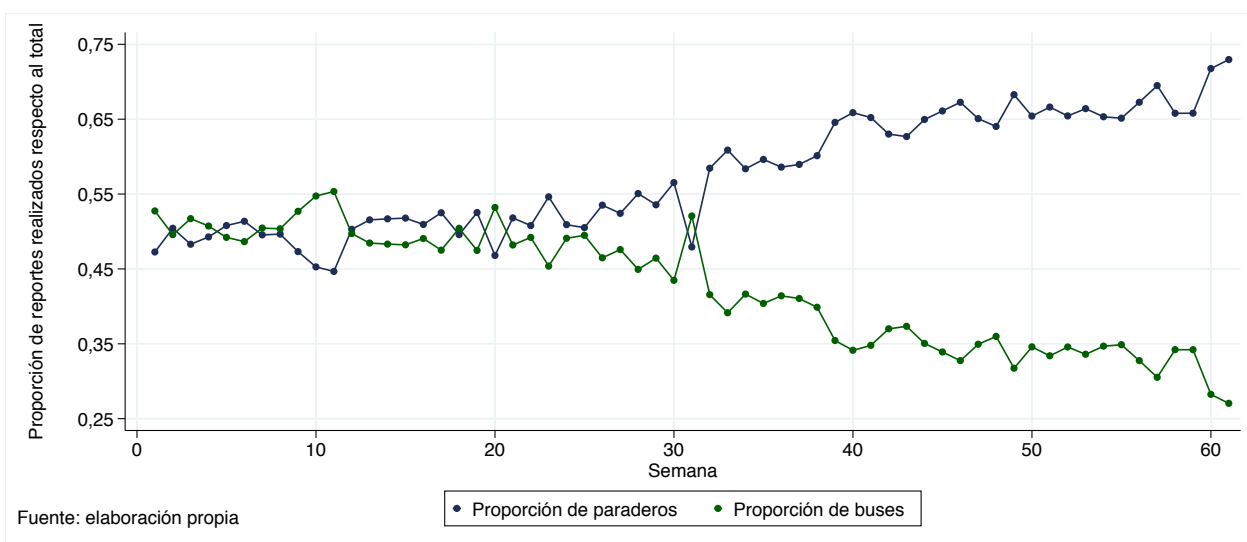
Ilustración 14: Reportes de paraderos y buses semanalmente.



En general se tiene que la cantidad de reportes realizados sobre buses y paraderos es bastante similar hasta la semana 32 (comienzo de Marzo 2019), donde la cantidad de reportes de paraderos se incrementa fuertemente hasta generar una diferencia que se sostiene en el tiempo (la cual se presume es debido al recambio de usuarios generado en Marzo y al aumento en su cantidad, potencialmente causado por publicidad en paraderos de la ciudad).

En la Ilustración 15 se grafica la proporción de reportes de buses y paraderos respecto al total de reportes realizados en la semana, donde se mantiene el mismo patrón: estabilidad alrededor del 50% de los reportes hasta la semana 32 en donde la proporción predominante es la de los reportes acerca de los paraderos, la cual incluso mantiene una tendencia positiva desde la semana 25.

Ilustración 15: Proporción de reportes de buses y paraderos respecto al total realizado, semanalmente.



3.3.2.2 Reportes positivos o pesimistas

En este subcapítulo se ve analiza descriptivamente los reportes diferenciando si son positivos o pesimistas¹⁴. En la Tabla 14, se indica la clasificación de cuáles reportes son positivos y cuáles pesimistas.

Tabla 14: Clasificación de reportes positivos y pesimistas en reportes de buses y paraderos.

| Buses | | Paradero | |
|-----------------------------|---------------|---------------------------------|---------------|
| Reporte | Clasificación | Reporte | Clasificación |
| Hay asientos disponibles | Positivo | Paradero sin problemas | Positivo |
| No hay asientos disponibles | Pesimista | Paradero lleno | Pesimista |
| Imposible moverse | Pesimista | No hay luz | Pesimista |
| Bus en buen estado | Positivo | No hay asiento | Pesimista |
| Sucio | Pesimista | No hay techo | Pesimista |
| Ventana en mal estado | Pesimista | No hay información de recorrido | Pesimista |
| Puerta en mal estado | Pesimista | 2 buses juntos | Pesimista |
| Pasamanos en mal estado | Pesimista | 3 o más buses juntos | Pesimista |
| Buen comportamiento | Positivo | Ebrio | Pesimista |
| Conducción brusca | Pesimista | Robos o desórdenes | Pesimista |
| No se detiene en paraderos | Pesimista | Bus en tránsito | Pesimista |
| Habla por celular | Pesimista | Bus vacío no se detuvo | Pesimista |
| Taco | Pesimista | Bus lleno no se detuvo | Pesimista |

Sólo hay 4 reportes positivos (3 de buses y 1 de paradero) que están asociados al buen estado de la infraestructura y buen comportamiento del chofer. El resto de los reportes son pesimistas, ya que corresponden a denuncias de mal estado de infraestructura, eventos que generan inseguridad, deficiente gestión de buses, y mala conducta del conductor.

También se considera la opción de la negación y confirmación del reporte, por lo que un reporte de categoría pesimista que es negado, se considera como un reporte positivo. Recíprocamente un reporte de categoría positiva que es negado, se considera como un reporte pesimista.

En la Tabla 15 se adjunta la cantidad de reportes mensuales realizados, cuántos son positivos y pesimistas, seguido de su proporción respectiva del total semanal.

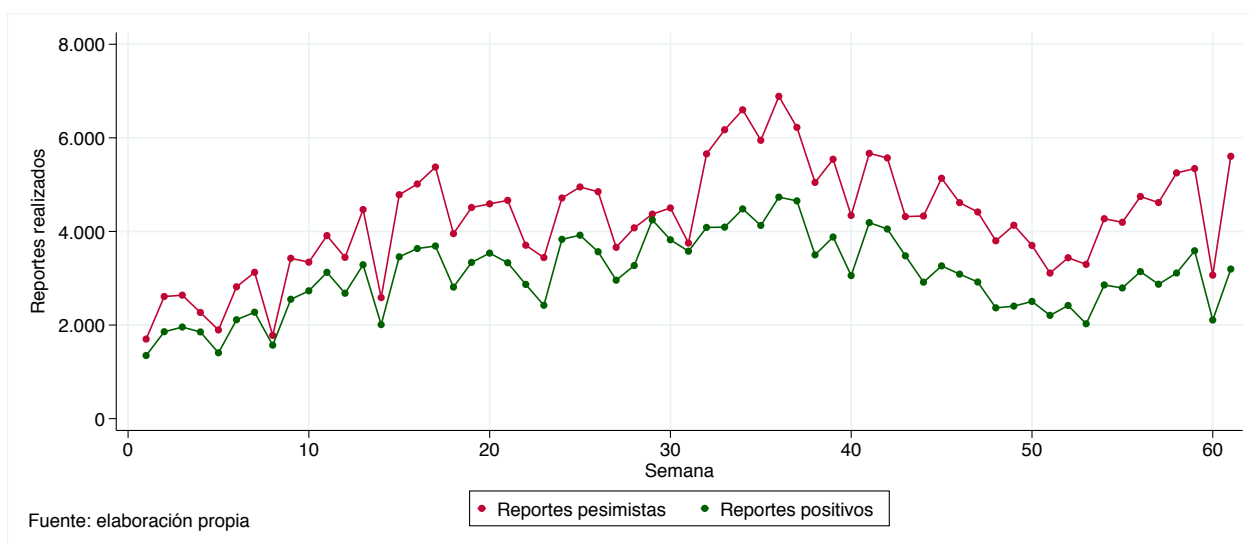
¹⁴ Un nombre más ideal sería negativo, pero se decide llamarle pesimista ya que también existen los reportes asociados a la negación de otros y debido a la no existencia de una palabra sinónima de negativo, se reserva para la otra categoría.

Tabla 15: Cantidad de reportes mensuales, acerca de paradero y bus.

| Año | Mes | Reportes realizados | Reportes pesimistas | Proporción de pesimistas | Reportes positivos | Proporción de positivos |
|------|------------|---------------------|---------------------|--------------------------|--------------------|-------------------------|
| 2018 | Agosto | 18.775 | 10.690 | 56,94% | 8.085 | 43,06% |
| | Septiembre | 20.402 | 11.565 | 56,69% | 8.837 | 43,31% |
| | Octubre | 29.236 | 16.523 | 56,52% | 12.713 | 43,48% |
| | Noviembre | 33.736 | 19.580 | 58,04% | 14.156 | 41,96% |
| | Diciembre | 31.870 | 18.245 | 57,25% | 13.625 | 42,75% |
| 2019 | Enero | 35.498 | 20.101 | 56,63% | 15.397 | 43,37% |
| | Febrero | 31.410 | 16.582 | 52,79% | 14.828 | 47,21% |
| | Marzo | 44.179 | 26.007 | 58,87% | 18.172 | 41,13% |
| | Abril | 42.009 | 24.568 | 58,48% | 17.441 | 41,52% |
| | Mayo | 38.984 | 22.582 | 57,93% | 16.402 | 42,07% |
| | Junio | 30.993 | 18.741 | 60,47% | 12.252 | 39,53% |
| | Julio | 25.820 | 15.555 | 60,24% | 10.265 | 39,76% |
| | Agosto | 32.232 | 19.553 | 60,66% | 12.679 | 39,34% |
| | Septiembre | 31.939 | 19.667 | 61,58% | 12.272 | 38,42% |

Se procede a realizar gráficos a partir de la variación de reportes positivos y pesimistas, utilizando un intervalo semanal (ver Ilustración 16 e Ilustración 17). La tabla de la cual se extrae la información se encuentra en Anexos, Apéndice D: Estadísticas de reportes positivos y pesimistas, según semana.

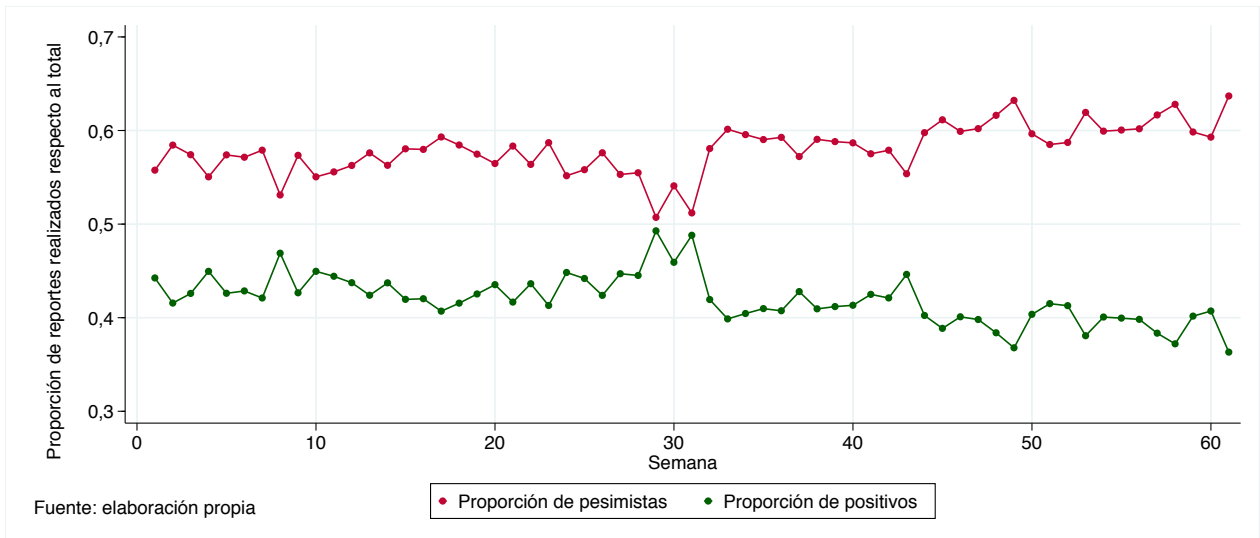
Ilustración 16: Reportes semanales positivos y pesimistas.



De la Ilustración 16 se observa como la cantidad de reportes pesimistas es constantemente mayor que los positivos. Al considerar la proporción que éstos representan de los reportes realizados

semanales se obtiene la Ilustración 17, en donde se observa cierta estabilidad. La proporción de reportes pesimistas fluctúa entre 50 y 60%, y la de los positivos entre 40 y 50%.

Ilustración 17: Proporción de reportes positivos y pesimistas respecto al total realizado, semanalmente.



3.3.2.3 Reportes afirmativos o negativos

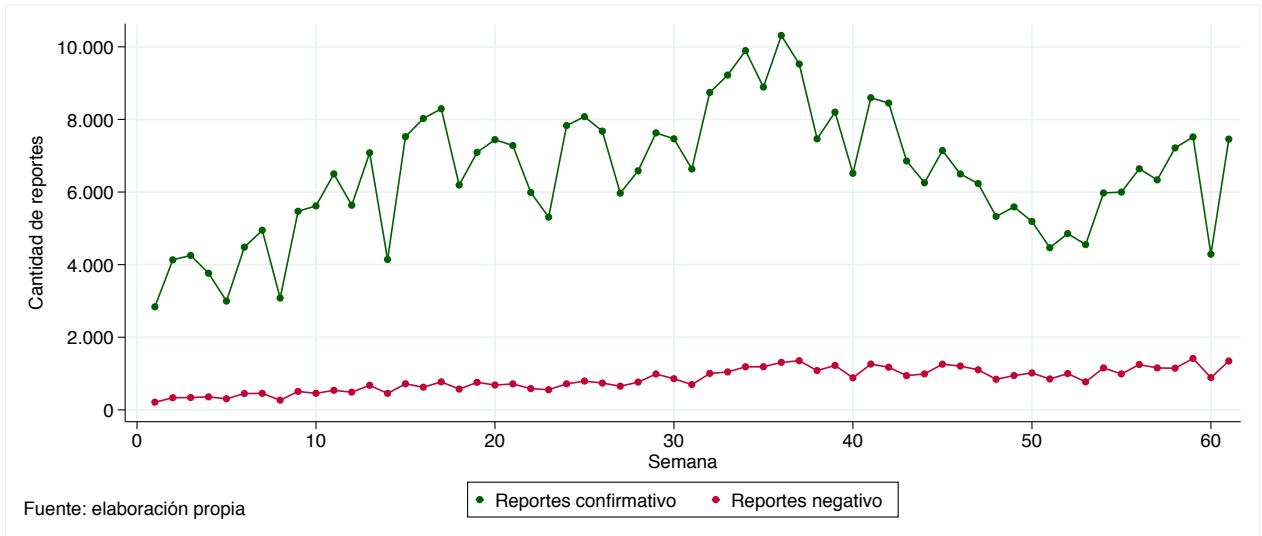
En la Tabla 16 se adjunta la cantidad de reportes mensuales realizados y cuántos son afirmativos y negativos, además de su respectiva proporción del total mensual.

Tabla 16: Cantidad de reportes confirmativos y negativos, semanalmente.

| Año | Mes | Reportes realizados | Reportes confirmativo | Proporción de confirmativos | Reportes negativos | Proporción de negativos |
|------|------------|---------------------|-----------------------|-----------------------------|--------------------|-------------------------|
| 2018 | Agosto | 18.775 | 17.339 | 48,94% | 1.436 | 51,06% |
| | Septiembre | 20.402 | 18.618 | 49,62% | 1.784 | 50,38% |
| | Octubre | 29.236 | 26.880 | 48,37% | 2.356 | 51,63% |
| | Noviembre | 33.736 | 30.930 | 51,13% | 2.806 | 48,87% |
| | Diciembre | 31.870 | 29.003 | 50,66% | 2.867 | 49,34% |
| 2019 | Enero | 35.498 | 32.365 | 51,99% | 3.133 | 48,01% |
| | Febrero | 31.410 | 28.087 | 53,81% | 3.323 | 46,19% |
| | Marzo | 44.179 | 39.466 | 58,72% | 4.713 | 41,28% |
| | Abril | 42.009 | 36.852 | 60,39% | 5.157 | 39,61% |
| | Mayo | 38.984 | 34.138 | 64,39% | 4.846 | 35,61% |
| | Junio | 30.993 | 26.380 | 65,75% | 4.613 | 34,25% |
| | Julio | 25.820 | 21.693 | 66,59% | 4.127 | 33,41% |
| | Agosto | 32.232 | 27.364 | 66,57% | 4.868 | 33,43% |
| | Septiembre | 31.939 | 27.019 | 68,89% | 4.920 | 31,11% |

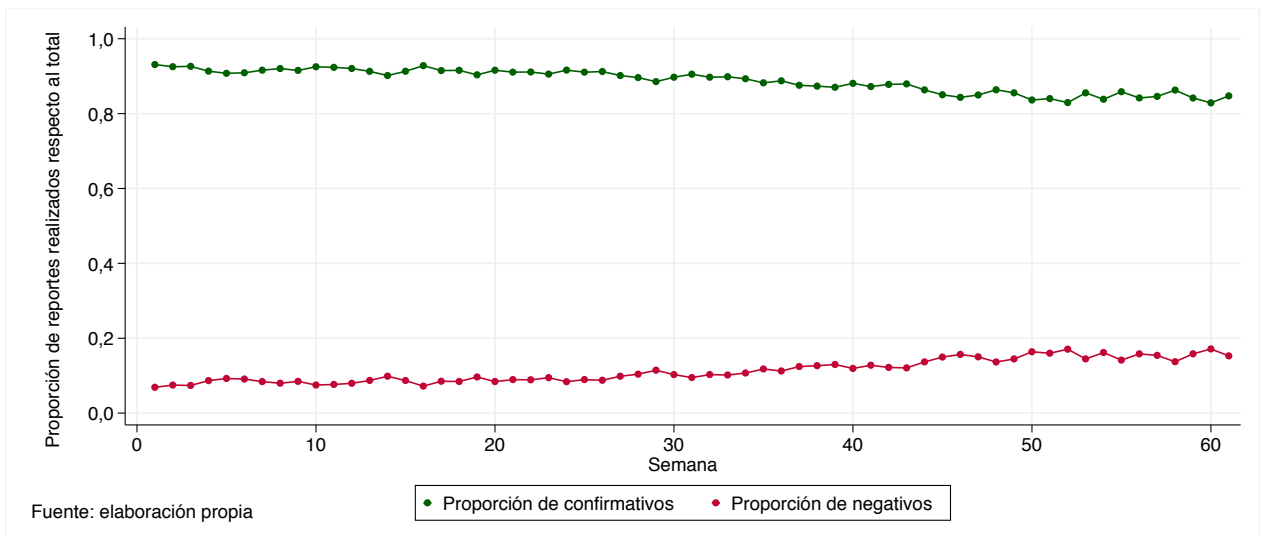
A continuación se graficará esta información (en la Ilustración 18 e Ilustración 19) pero utilizando el período semanal, cuya tabla (símil a la previa) está en Anexos, Apéndice E: Estadísticas de reportes confirmativos y negativos, según semana.

Ilustración 18: Reportes confirmativos y negativos semanales.



En la Ilustración 18 se ve cómo la cantidad de reportes afirmativos mantiene un patrón similar a los reportes totales (ver Ilustración 11), a diferencia de los reportes declinativos que mantiene una estabilidad creciente. Al graficar la proporción de éstos tipos de reportes se obtiene el siguiente gráfico (ver Ilustración 19), donde se muestra claramente cómo cada vez hay más reportes declinativos y menos confirmativos, proporcionalmente.

Ilustración 19: Proporción semanal de reportes confirmativos y negativos respecto al total realizado.



3.3.2.4 Reportes según la hora del día

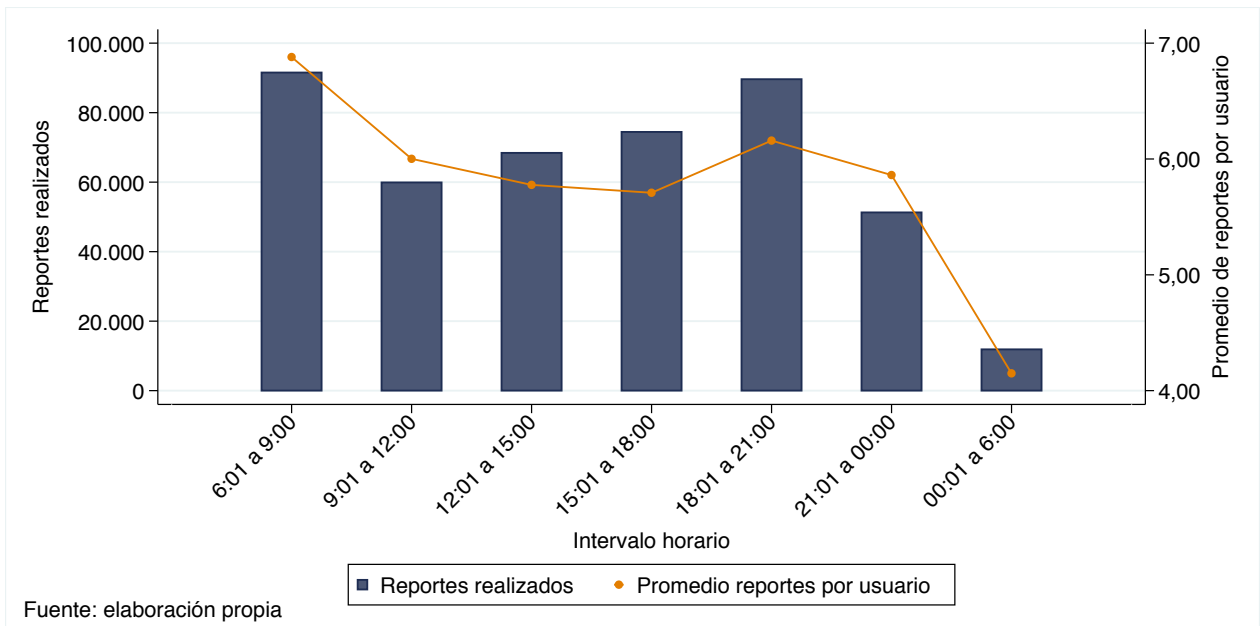
Se procede a realizar el análisis de los reportes realizados según la hora del día, para esto se trabaja con los intervalos previamente establecidos. En la Tabla 17 se adjunta la frecuencia de reportes realizados según intervalo horario.

Tabla 17: Cantidad de reportes y usuarios reportadores, según intervalo horario.

| Intervalo horario | Reportes realizados | Porcentaje | Acumulado | Usuarios reportadores | Promedio reportes por usuario |
|-------------------|---------------------|----------------|-----------|-----------------------|-------------------------------|
| 6:01 a 9:00 | 91.540 | 20,47% | 20,47% | 13.305 | 6,88 |
| 9:01 a 12:00 | 59.913 | 13,40% | 33,88% | 9.983 | 6,00 |
| 12:01 a 15:00 | 68.415 | 15,30% | 49,18% | 11.844 | 5,78 |
| 15:01 a 18:00 | 74.450 | 16,65% | 65,83% | 13.042 | 5,71 |
| 18:01 a 21:00 | 89.615 | 20,04% | 85,88% | 14.551 | 6,16 |
| 21:01 a 00:00 | 51.280 | 11,47% | 97,35% | 8.749 | 5,86 |
| 00:01 a 6:00 | 11.870 | 2,65% | 100,00% | 2.861 | 4,15 |
| Total | 447.083 | 100,00% | - | - | - |

En la Ilustración 20 se visualiza la información de la tabla previa, en donde los peaks en cantidad de reportes realizados son en las horas del día donde más viajes se realizan en la ciudad, las horas peak de la mañana y la tarde. No obstante, la cantidad de reportes por usuario, en promedio, es mayor en la mañana que en la tarde.

Ilustración 20: Reportes realizados y promedio de reportes por usuario según intervalo horario.



3.3.2.5 Reportes según día de la semana

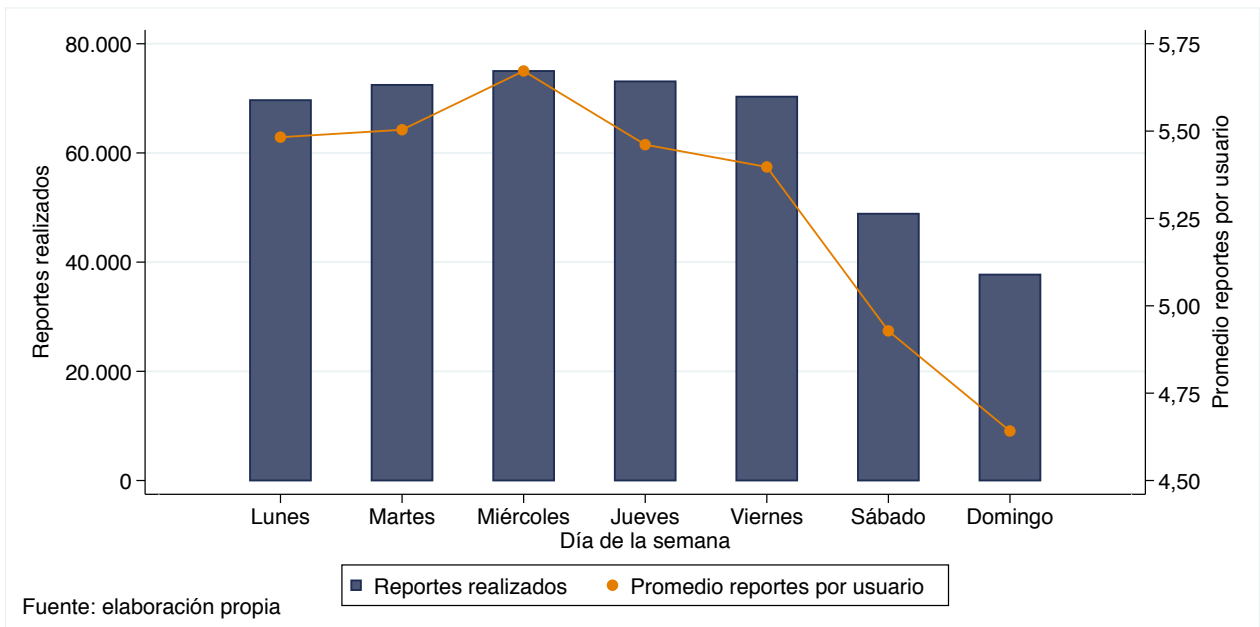
En este subcapítulo se realiza el mismo análisis descriptivo del subcapítulo previo, pero ahora según el día de la semana. Se tiene la Tabla 18 con la cantidad de reportes y el promedio de reportes por usuario.

Tabla 18: Cantidad de reportes y usuarios reportadores, según día de la semana.

| Día de la semana | Reportes realizados | Porcentaje | Acumulado | Usuarios reportadores | Promedio reportes por usuario |
|------------------|---------------------|----------------|-----------|-----------------------|-------------------------------|
| Lunes | 69.660 | 15,58% | 15,58% | 12.706 | 5,48 |
| Martes | 72.452 | 16,21% | 31,79% | 13.164 | 5,50 |
| Miércoles | 75.000 | 16,78% | 48,56% | 13.223 | 5,67 |
| Jueves | 73.102 | 16,35% | 64,91% | 13.386 | 5,46 |
| Viernes | 70.300 | 15,72% | 80,64% | 13.025 | 5,40 |
| Sábado | 48.857 | 10,93% | 91,56% | 9.914 | 4,93 |
| Domingo | 37.712 | 8,44% | 100,00% | 8.125 | 4,64 |
| Total | 447.083 | 100,00% | - | - | - |

Luego, se grafica esta información y se observa que la cantidad de reportes es estable durante los días laborales y disminuye en los fines de semana (ver Ilustración 21). El promedio de reportes por usuario tiene un patrón similar al descrito, no obstante, es interesante el caso del día domingo donde la cantidad de reportes realizados se reduce cerca de un 50%, pero la cantidad de usuarios que reportan se reduce sólo en 38%.

Ilustración 21: Reportes realizados y promedio de reportes por usuario según día de la semana.



3.3.3 Problema de participación

Tal como se indica en capítulo “Participación en plataformas online” de la revisión bibliográfica, en éstas existe un problema de participación en los usuarios que sigue un patrón de 90:9:1, en donde el 90% de los usuarios no aporta contenido, el 9% lo hace pero en baja cantidad y el 1% restante de los usuarios realiza gran parte de las contribuciones. En esta sección se corrobora y profundiza si esta regla está presente en *Transapp*.

En la Tabla 20 (próxima página) se adjuntan diferentes variables importantes que permiten analizar la existencia de esta problemática. Se tiene en primer lugar la cantidad de usuarios consultores (quienes utilizaron efectivamente la aplicación) en cada mes y la cantidad de usuarios reportadores; con éstas se calcula la proporción de usuarios que reportan en el mes. Posteriormente, se calcula el 1% de los usuarios consultores y se obtiene la suma de cuántos reportes realizaron el 1% de los usuarios con mayor número de reportes. Finalmente, con ésta variable se estima la proporción que representa respecto al número total de reportes mensuales, y así es posible ver la densidad de reportes que acumulan los usuarios reportadores con “mejor desempeño”.

En la Tabla 19, se realiza un contraste de la Regla de Nielsen, utilizando la información de la Tabla 20 (que considera todos los reportes de usuarios, sin importar su clasificación).

Tabla 19: Contraste de reportes en *Transapp* con la Regla de Nielsen.

| Regla de Nielsen Porcentaje de usuarios | 90% | 9% | 1% |
|---|--------|--------|------------------|
| Porcentaje promedio en <i>Transapp</i> | 83,27% | 15,73% | 1% ¹⁵ |
| Porcentaje promedio de reportes del total | 0% | 51,68% | 48,32% |

Se tiene que, utilizando el promedio mensual, el 83,27% de los usuarios consultores no reporta y el 16,73% usuarios sí reporta. Al segmentar éstos últimos, se tiene que el 15,73% de los usuarios realiza el 51,68% de los reportes, y el 1% restante realiza el 48,32% de los reportes.

Se concluye con la afirmación de que en *Transapp* la Regla de Nielsen sí está presente, pues se cumple con sus principales enunciados, gran cantidad de usuarios que no contribuyen, pocos usuarios que contribuyen levemente, y una reducida proporción que contribuye de gran manera.

¹⁵ Se construye la tabla de forma tal de fijar el 1% de usuarios.

Tabla 20: Estadísticas asociadas a la Regla de Nielsen en Transapp. Cantidad de reportes, reportadores y proporciones de reportes según la clasificación de usuario.

| Año | Mes | Usuarios consultores | Usuarios reportadores | Reportes realizados | Porcentaje de reportadores | 1% de los usuarios consultores | Reportes realizados por el 1% | Porcentaje de reportes del total |
|------|------------|----------------------|-----------------------|---------------------|----------------------------|--------------------------------|-------------------------------|----------------------------------|
| 2018 | Agosto | 13.504 | 2.535 | 18.775 | 18,77% | 135 | 8.665 | 46,15% |
| | Septiembre | 17.303 | 3.270 | 20.402 | 18,90% | 173 | 8.092 | 39,66% |
| | Octubre | 20.733 | 3.960 | 29.236 | 19,10% | 207 | 13.348 | 45,66% |
| | Noviembre | 23.459 | 4.687 | 33.736 | 19,98% | 234 | 15.386 | 45,61% |
| | Diciembre | 26.375 | 4.868 | 31.870 | 18,46% | 263 | 14.409 | 45,21% |
| 2019 | Enero | 28.478 | 5.325 | 35.498 | 18,70% | 284 | 17.622 | 49,64% |
| | Febrero | 26.937 | 4.578 | 31.410 | 17,00% | 269 | 16.402 | 52,22% |
| | Marzo | 36.275 | 6.772 | 44.179 | 18,67% | 362 | 20.416 | 46,21% |
| | Abril | 37.938 | 6.601 | 42.009 | 17,40% | 379 | 19.723 | 46,95% |
| | Mayo | 37.474 | 6.043 | 38.984 | 16,13% | 374 | 19.532 | 50,10% |
| | Junio | 37.181 | 5.181 | 30.993 | 13,93% | 371 | 15.412 | 49,73% |
| | Julio | 37.708 | 4.650 | 25.820 | 12,33% | 377 | 13.246 | 51,30% |
| | Agosto | 40.637 | 5.298 | 32.232 | 13,04% | 406 | 17.449 | 54,14% |
| | Septiembre | 47.320 | 5.601 | 31.939 | 11,84% | 473 | 17.235 | 53,96% |

3.4. Análisis de viajes y evaluaciones

Antes de iniciar con el trabajo asociado a la base de datos de los viajes y evaluaciones de éstos realizadas por los usuarios, hay que realizar un filtro para así cerciorarse de que efectivamente corresponden a viajes reales y no sean “ficticios” o erróneos. Para esto, se conserva los viajes que tienen una distancia euclidiana mayor a 500 metros (entre el inicio y el final del recorrido) y menor a 48 kilómetros (el recorrido del bus más largo de Santiago (Correa, 2017)), una duración del viaje superior a 3 minutos y menor a 175 minutos¹⁶. Los viajes que no cumplen estas condiciones, se eliminaron de la base de datos.

La base original contiene 181.847 viajes entre Agosto de 2018 a Septiembre de 2019 y con el filtro se elimina 89.600 viajes equivalentes al 49,43% de los datos, por lo que la base final con la que se trabaja contiene 91.961 viajes, equivalente al 50,57% de los datos originales.

3.4.1 Análisis descriptivo

3.4.1.1 Estadísticas básicas de viajes y evaluaciones

En la Tabla 21 (próxima página) se adjunta la distribución mensual de los viajes realizados, la cantidad de usuarios viajeros, el promedio de viajes por usuario; y de igual manera para las evaluaciones, se presenta la cantidad, el número de usuarios evaluadores, y el promedio de evaluaciones por usuario (que evalúa). Finalmente, se tiene el porcentaje de la cantidad de viajes mensuales realizados que son evaluados, y la proporción de los usuarios viajeros que evalúa viajes.

Si bien la siguiente tabla está en formato mensual, los siguientes gráficos del capítulo se realizan utilizando una distribución semanal, por lo que en el Apéndice F: Estadísticas de viajes y evaluaciones semanales, se encuentra una tabla idéntica pero utilizando aquella unidad temporal. La razón es igual a la de los subcapítulos previos, normalizar la cantidad de días laborales y no-laborales.

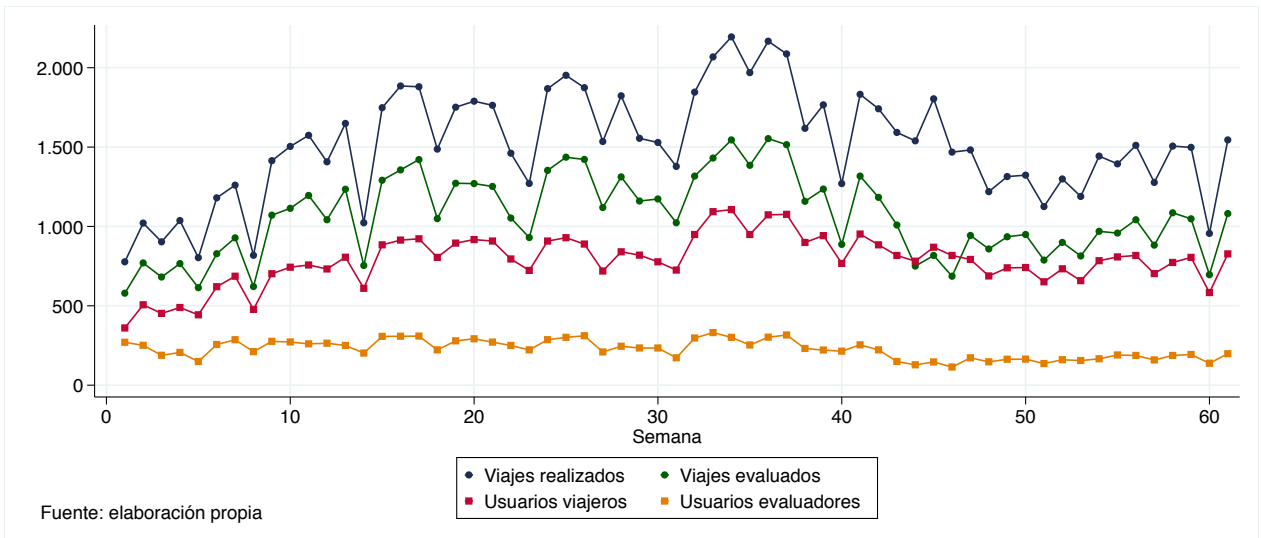
¹⁶ Si bien en la base de datos existían viajes de 180 minutos, había una anomalía en donde se acumulaban muchos en esta cercanía, por lo que se eliminaron debido a que se cree que es una respuesta automática de la aplicación a terminar viajes excesivamente largos (o que el usuario no finaliza).

Tabla 21: Cantidad de viajes, evaluaciones y usuarios viajeros y evaluadores, mensualmente.

| Año | Mes | Viajes realizados | Usuarios que viajan | Promedio viajes por usuario | Evaluaciones realizadas | Usuarios que evalúan | Promedio evaluaciones por usuario | Porcentaje viajes evaluados | Porcentaje usuarios que evalúan |
|------|------------|-------------------|---------------------|-----------------------------|-------------------------|----------------------|-----------------------------------|-----------------------------|---------------------------------|
| 2018 | Agosto | 4.362 | 1.377 | 3,17 | 3.286 | 1.017 | 3,23 | 75,33% | 73,86% |
| | Septiembre | 4.851 | 1.861 | 2,61 | 3.572 | 1.076 | 3,32 | 73,63% | 57,82% |
| | Octubre | 6.652 | 2.252 | 2,95 | 4.962 | 1.133 | 4,38 | 74,59% | 50,31% |
| | Noviembre | 7.207 | 2.468 | 2,92 | 5.288 | 1.218 | 4,34 | 73,37% | 49,35% |
| | Diciembre | 7.062 | 2.562 | 2,76 | 5.053 | 1.135 | 4,45 | 71,55% | 44,30% |
| 2019 | Enero | 7.952 | 2.692 | 2,95 | 5.861 | 1.244 | 4,71 | 73,70% | 46,21% |
| | Febrero | 6.224 | 2.208 | 2,82 | 4.635 | 896 | 5,17 | 74,47% | 40,58% |
| | Marzo | 8.686 | 3.019 | 2,88 | 6.110 | 1.256 | 4,86 | 70,34% | 41,60% |
| | Abril | 7.938 | 2.876 | 2,76 | 5.662 | 1.123 | 5,04 | 71,33% | 39,05% |
| | Mayo | 7.362 | 2.590 | 2,84 | 4.813 | 882 | 5,46 | 65,38% | 34,05% |
| | Junio | 6.285 | 2.284 | 2,75 | 3.436 | 611 | 5,62 | 54,67% | 26,75% |
| | Julio | 5.502 | 2.132 | 2,58 | 3.870 | 678 | 5,71 | 70,34% | 31,80% |
| | Agosto | 6.270 | 2.344 | 2,67 | 4.293 | 779 | 5,51 | 68,47% | 33,23% |
| | Septiembre | 5.608 | 2.178 | 2,57 | 3.984 | 739 | 5,39 | 71,04% | 33,93% |

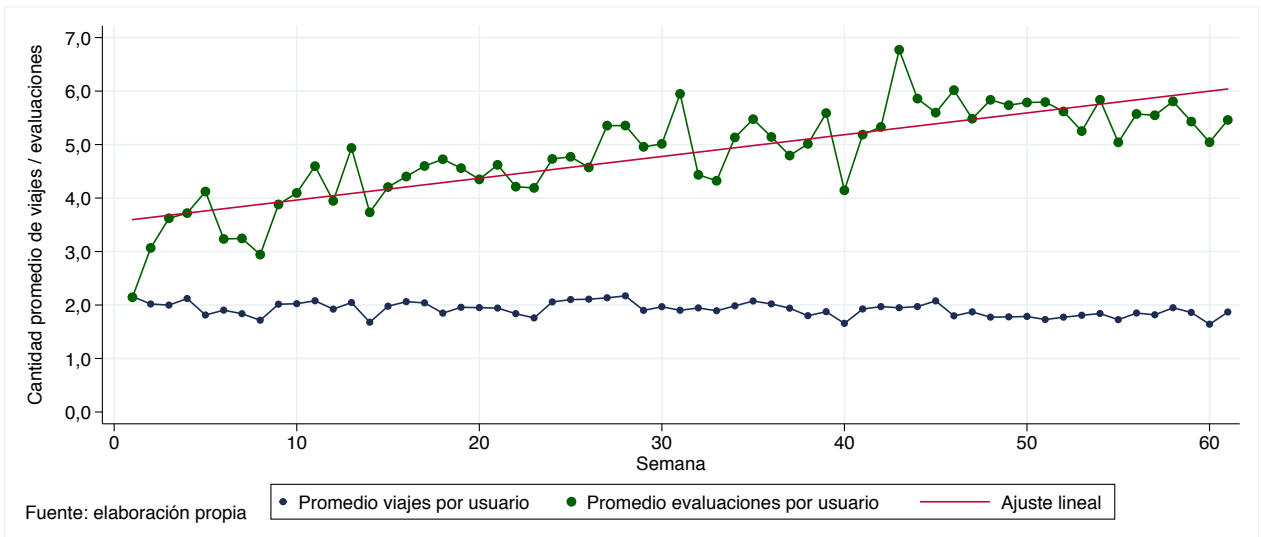
En conjunto con la información de la Tabla 21 y la Ilustración 22, en general se tiene que los viajes realizados, viajes evaluados, usuarios viajeros y evaluadores tienen patrones similares en el tiempo.

Ilustración 22: Cantidad de viajes, evaluaciones realizadas, usuarios viajeros y usuarios evaluadores, semanalmente.



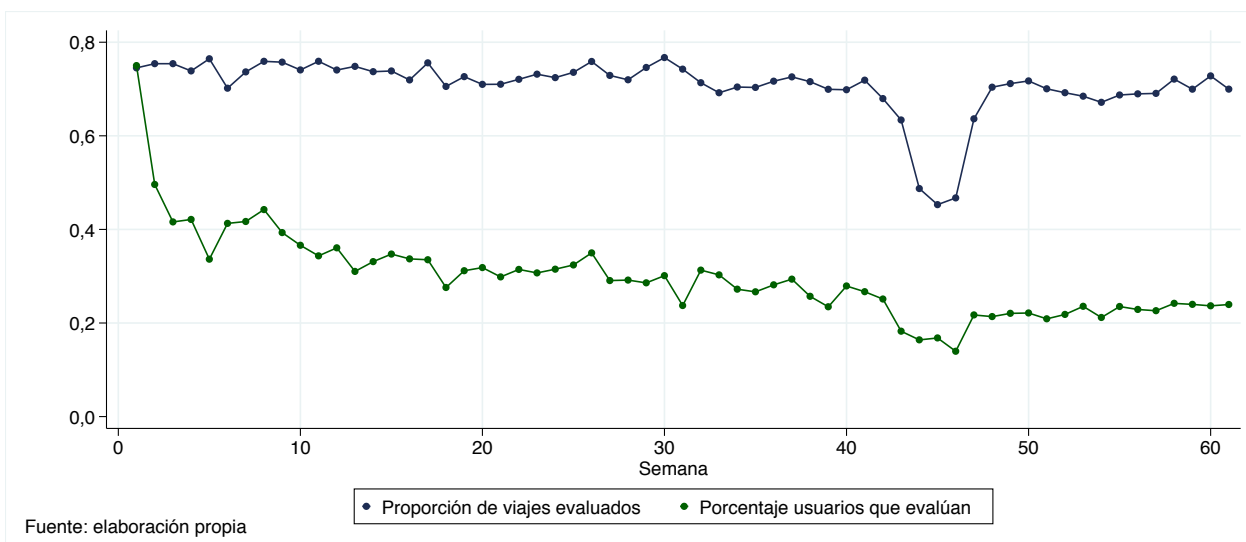
En la Ilustración 23, se observa la cantidad promedio de viajes por usuario, la cual se mantiene estable en torno a los 2 viajes. Por el otro lado, la cantidad promedio de evaluaciones por usuario (que evalúa) mantiene una tendencia positiva, siendo cercana a los 3 viajes evaluados en un comienzo en la semana 1, a más de 5 desde la semana 30, la cual se observa en el ajuste lineal.

Ilustración 23: Cantidad promedio de viajes y evaluaciones por usuario, semanalmente.



Al estudiar la proporción de viajes evaluados (ver Ilustración 24) en general hay un patrón estable, exceptuando una anomalía en la cercanía de la semana 45, durante las últimas dos semanas de Mayo y la primera de Junio de 2019. Se desconoce una potencial causa. Por el otro lado, la proporción de viajeros que realizan viajes y los evalúan presenta una disminución progresiva.

Ilustración 24: Proporción de viajes evaluados y proporción de viajeros que evalúan viajes, según semana.



Esta información sumada a la de la Ilustración 23, se concluye que si bien la proporción de usuarios evaluadores disminuye en el tiempo, quienes siguen haciéndolo, evalúan cada vez más. Esto es negativo ya que si se considera que los usuarios mantienen hábitos de viaje, hay recorridos que antes eran evaluados (por ciertos usuarios) y ya no, provocando así una potencial disminución en la cantidad de recorridos evaluados (ya que son una menor proporción quienes evalúan). Este análisis es profundizado en Anexos, Apéndice G: Estadísticas de recorridos, en donde se demuestra que efectivamente hay menos recorridos evaluados a medida que transcurre el tiempo.

3.4.1.2 Estadísticas de viajeros y evaluadores

Al considerar la cantidad de viajes realizados en total en el período se presenta la siguiente distribución (ver Tabla 22), en donde la mayoría de los usuarios (92%) tiene menos de 10 viajes realizados, y menos del 1% tiene más de 51 viajes.

Tabla 22: Cantidad y proporción de usuarios que realizan viajes en determinadas semanas.

| Cantidad de viajes | Cantidad de usuarios | Proporción de usuarios | Acumulado |
|-----------------------|----------------------|------------------------|-----------|
| Sólo 1 | 9.904 | 49,35% | 49,35% |
| Entre 2 y 4 | 6.091 | 30,35% | 79,71% |
| Entre 5 y 10 | 2.472 | 12,32% | 92,03% |
| Entre 11 y 25 | 1.089 | 5,43% | 97,45% |
| Entre 26 y 50 | 319 | 1,59% | 99,04% |
| Más de 51 | 192 | 0,96% | 100,00% |
| Total usuarios | 20.067 | 100,00% | - |

En la Ilustración 25, se grafica a la izquierda un histograma de usuarios según los viajes realizados, sólo para aquellos que viajan en menos de 25 oportunidades (97,5% de los usuarios). Se omite un análisis de logaritmo (como en los subcapítulos de consultas y reportes, ya que el máximo de viajes realizados por un usuario es de 391).

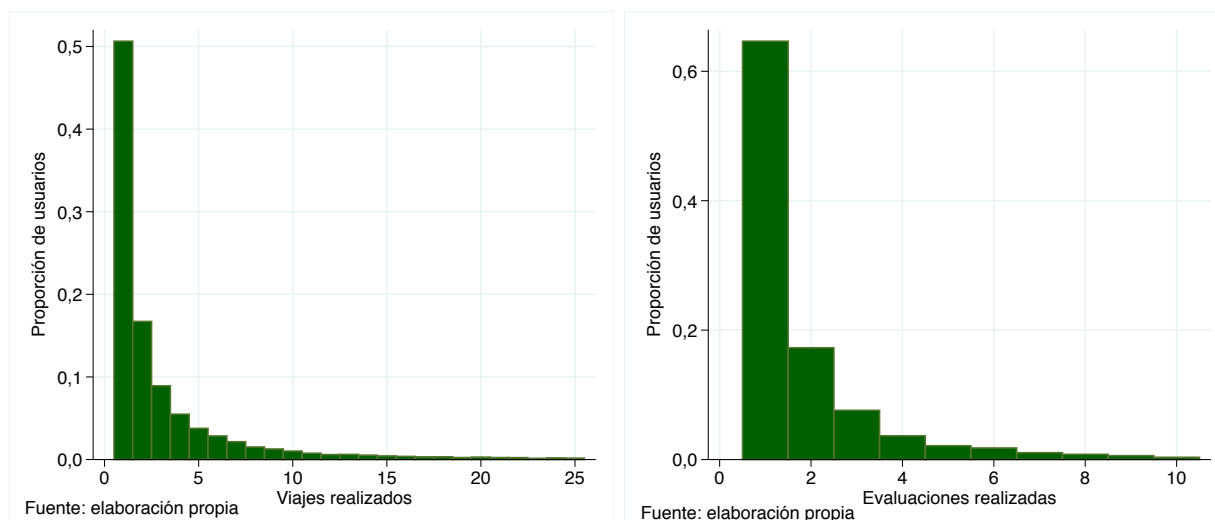
Por otro lado, si se considera la cantidad de viajes evaluados la situación es idéntica, el 94% de los usuarios evalúa hasta 10 viajes, y menos del 1% evalúa más de 51 viajes.

Tabla 23: Cantidad y proporción de usuarios que realizan evaluaciones de viajes en determinadas semanas.

| Cantidad de viajes evaluados | Cantidad de usuarios | Proporción de usuarios | Acumulado |
|------------------------------|----------------------|------------------------|-----------|
| Sólo 1 | 8.163 | 59,21% | 59,21% |
| Entre 2 y 4 | 3.627 | 26,31% | 85,52% |
| Entre 5 y 10 | 1.177 | 8,54% | 94,05% |
| Entre 11 y 25 | 537 | 3,89% | 97,95% |
| Entre 26 y 50 | 172 | 1,25% | 99,19% |
| Más de 51 | 111 | 0,81% | 100,00% |
| Total usuarios | 13.787 | 100,00% | - |

En el grafico de la derecha de la Ilustración 25, se tiene un histograma de usuarios según cantidad de evaluaciones realizadas, sólo para quienes evalúan menos de 10 veces (94% de los usuarios). Se omite un análisis de logaritmo, ya que la cantidad de viajes máximos realizados por un usuario es 304.

Ilustración 25: Proporción de usuarios según viajes realizados (izquierda), y según evaluaciones realizadas (derecha).



3.4.2 Análisis de viajes y evaluaciones locales

En este apartado, se procede a analizar cómo varía los viajes y evaluaciones según la hora del día y el día de la semana.

3.4.2.1 Análisis según horario

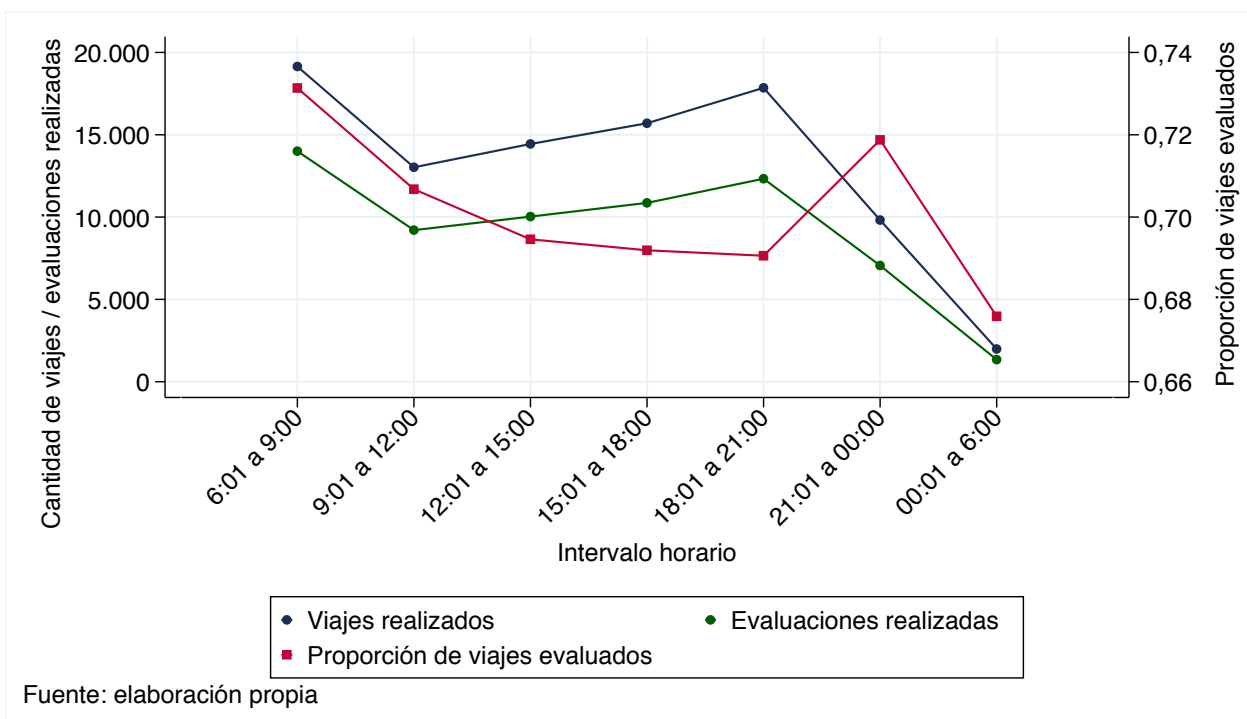
En la Tabla 24 se presenta la cantidad de viajes y evaluaciones realizadas según la hora del día, en intervalos de cada 3 horas a excepción de la noche-madrugada en donde se aumenta a un intervalo de 6 horas, esto debido a la baja de viajeros en aquel período.

Tabla 24: Frecuencia de viajes y evaluaciones realizadas según hora del día.

| Intervalo horario | Viajes | | | Evaluaciones | | |
|-------------------|-------------------|------------|-----------|-------------------------|------------|-----------|
| | Viajes realizados | Porcentaje | Acumulado | Evaluaciones realizadas | Porcentaje | Acumulado |
| 6:01 a 9:00 | 19.144 | 20,82% | 20,82% | 14.001 | 21,60% | 21,60% |
| 9:01 a 12:00 | 13.024 | 14,16% | 34,98% | 9.205 | 14,20% | 35,80% |
| 12:01 a 15:00 | 14.440 | 15,70% | 50,68% | 10.030 | 15,47% | 51,27% |
| 15:01 a 18:00 | 15.698 | 17,07% | 67,75% | 10.862 | 16,76% | 68,03% |
| 18:01 a 21:00 | 17.851 | 19,41% | 87,16% | 12.328 | 19,02% | 87,04% |
| 21:01 a 00:00 | 9.817 | 10,68% | 97,84% | 7.056 | 10,88% | 97,93% |
| 00:01 a 6:00 | 1.987 | 2,16% | 100,00% | 1.343 | 2,07% | 100,00% |
| Total | 91.961 | 100,00% | - | 64.825 | 100,00% | - |

Esta información se grafica en la siguiente Ilustración 26, en donde se observa que hay dos peaks, tanto en la cantidad de viajes como de evaluaciones realizadas. Estas son ambas en la hora punta mañana y tarde, lo cual es esperable dado que son los intervalos con mayor cantidad de viajes en la ciudad. Por otro lado, al estudiar la proporción de viajes evaluados según la hora del día, ésta fluctúa entre 73% y 67%. En general decrece a lo largo del día, exceptuando el período de 21 hrs. a 24 hrs. en donde aumenta a su segundo mayor valor.

Ilustración 26: Consultas, evaluaciones y proporción de viajes evaluados, según hora del día.



Al analizar si la diferencia en la proporción de viajes evaluados es estadísticamente diferente, se realiza una regresión con variable dependiente la proporción de viajes evaluados, y variables independientes los diferentes intervalos horarios en formato binario (=1 si lo es, y 0 si no), en donde el grupo base es el intervalo de 18:01 a 21:00 hrs (debido a que es el con menor proporción), cuyos resultados se adjuntan en la Tabla 25.

Ecuación 1: Especificación modelo para testear diferencias en la probabilidad de evaluar el viaje según intervalo horario.

$$\text{Viaje evaluado}_i = B_0 + \sum_{i=1}^6 B_i * \text{Intervalo horario}_i + \varepsilon$$

Tabla 25: Resultados regresión de proporción de viajes evaluados según día de la semana.

| Coefficiente | Estimador | Err. estándar | Valor t | Pr(> t) |
|---------------|-----------|---------------|---------|----------|
| 6:01 a 9:00 | 0,0407 | 0,004 | 8,59 | 0,000 |
| 9:01 a 12:00 | 0,0162 | 0,005 | 3,08 | 0,002 |
| 12:01 a 15:00 | 0,0040 | 0,005 | 0,78 | 0,434 |
| 15:01 a 18:00 | 0,0013 | 0,005 | 0,27 | 0,790 |
| 21:01 a 00:00 | 0,0281 | 0,006 | 4,91 | 0,000 |
| 00:01 a 6:00 | -0,01487 | 0,010 | -1,36 | 0,172 |
| Intercepto | 0,6901 | 0,003 | 202,44 | 0,000 |

Se concluye que la diferencia en la proporción es significativa para los intervalos de 6:01 a 9:00, 9:01 a 12:00, y 21:01 a 00:00. En el resto de los intervalos no hay evidencia estadística de diferencias significativas respecto a la proporción de las 18:01 a 21:00 hrs.

De lo anterior, se desprende que las personas son ligeramente más propensas a realizar evaluaciones de sus viajes durante la mañana y posterior a la hora punta de la tarde.

3.4.2.2 Análisis según día de la semana

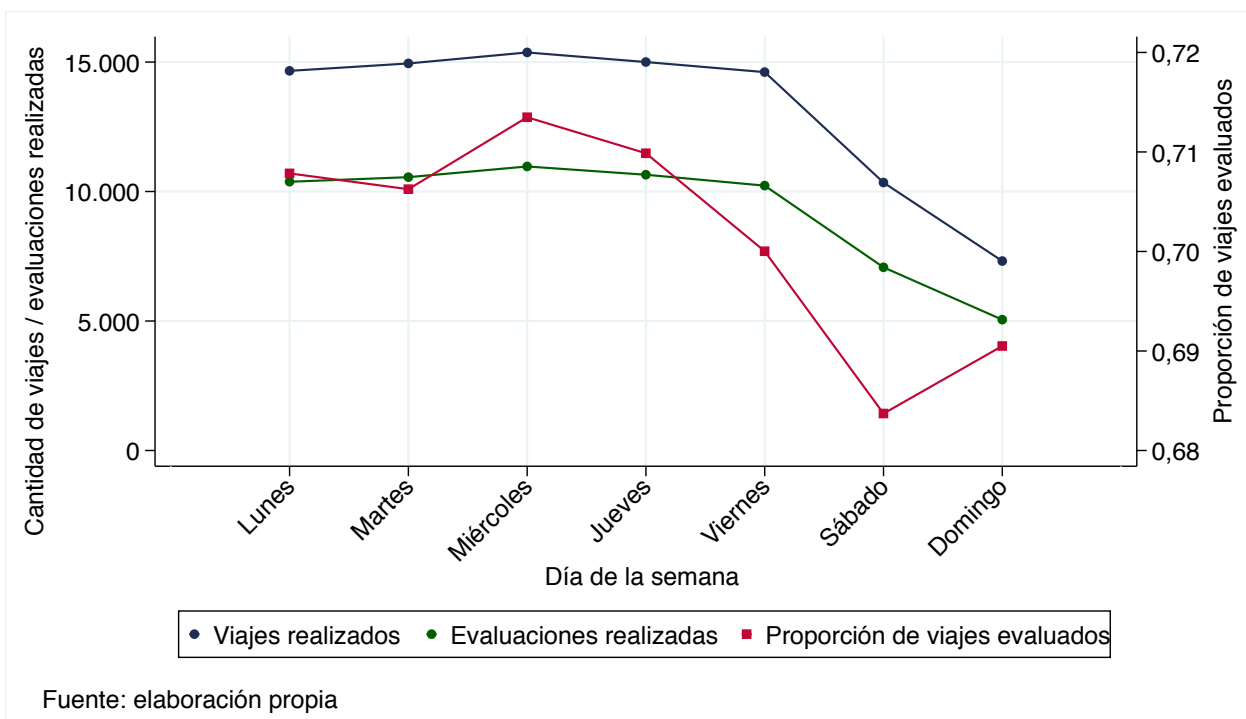
En este subcapítulo se procede a realizar el mismo análisis pero modificando según el día de la semana, cuyas estadísticas se encuentran en la Tabla 26. Se espera evidenciar diferencias entre los días lunes-jueves, viernes, sábado y domingo, debido a que los ciudadanos tienen patrones de viajes diferentes para cada uno de ellos.

Tabla 26: Frecuencia de viajes y evaluaciones realizadas según día de la semana.

| Intervalo horario | Viajes | | | Evaluaciones | | |
|-------------------|-------------------|----------------|-----------|-------------------------|------------|-----------|
| | Viajes realizados | Porcentaje | Acumulado | Evaluaciones realizadas | Porcentaje | Acumulado |
| Lunes | 14.612 | 15,89% | 15,89% | 10.365 | 15,99% | 15,99% |
| Martes | 14.899 | 16,20% | 32,09% | 10.542 | 16,26% | 32,25% |
| Miércoles | 15.329 | 16,67% | 48,76% | 10.956 | 16,90% | 49,15% |
| Jueves | 14.952 | 16,26% | 65,02% | 10.635 | 16,41% | 65,56% |
| Viernes | 14.557 | 15,83% | 80,85% | 10.218 | 15,76% | 81,32% |
| Sábado | 10.318 | 11,22% | 92,07% | 7.063 | 10,90% | 92,22% |
| Domingo | 7.294 | 7,93% | 100,00% | 5.046 | 7,78% | 100,00% |
| Total | 91.961 | 100,00% | - | 64.825 | 100,00% | - |

La cantidad de viajes es estable durante los días laborales, y en el fin de semana disminuyen fuertemente su número. En la Ilustración 27 se observa esta situación, y además se grafica la proporción de viajes evaluados.

Ilustración 27: Consultas, evaluaciones y proporción de viajes evaluados, según día de la semana.



La proporción de viajes evaluados según el día está comprendida entre 68% y 72%, presentando ligeras diferencias. Destaca principalmente que en los sábados y domingos hay menor proporción de viajes evaluados que en los días laborales, donde la proporción es relativamente estable. Para ver si estas diferencias son estadísticamente significativas, se realiza una regresión con variable dependiente la proporción de viajes evaluados y variables independientes los diferentes días de la semana, con sábado siendo el grupo de control, cuyos resultados se muestran en la Tabla 27.

Ecuación 2: Especificación del modelo para testear diferencias en la probabilidad de evaluar el viaje según el día de la semana.

$$\text{Viaje evaluado}_i = B_0 + \sum_{i=1}^6 B_i * \text{Día Semana}_i + \varepsilon$$

Tabla 27: Resultados regresión de proporción de viajes evaluados según día de la semana.

| Coeficiente | Estimador | Err. estándar | Valor t | Pr(> t) |
|-------------|-----------|---------------|---------|----------|
| Lunes | 0,0248 | 0,006 | 4,23 | 0,000 |
| Martes | 0,0230 | 0,006 | 3,94 | 0,000 |
| Miércoles | 0,0302 | 0,006 | 5,20 | 0,000 |
| Jueves | 0,0267 | 0,006 | 4,58 | 0,000 |
| Viernes | 0,0174 | 0,006 | 2,96 | 0,003 |
| Domingo | 0,0073 | 0,007 | 1,04 | 0,297 |
| Intercepto | 0,6845 | 0,004 | 152,48 | 0,000 |

Se tiene que la diferencia entre los días laborales y el sábado es significativa, y que la diferencia entre el sábado y domingo no lo es. Dado lo anterior, se concluye que las personas son ligeramente más propensas a realizar evaluaciones de sus viajes durante los días laborales que los fines de semana.

3.4.3 Análisis descriptivo de la nota del viaje

Se recuerda que al finalizar el viaje realizado en la aplicación, los usuarios son invitados a evaluarlo (pueden no hacerlo), para lo cual se les indica las posibles evaluaciones, las que han sido modificadas a través del tiempo, ya que previo a Mayo de 2019 se presentaban tres sapos¹⁷ de diferentes colores aludiendo a un semáforo (rojo, amarillo y verde); y posterior a Junio de 2019 (luego de un proceso de transición), se agregaron dos nuevas evaluaciones intermedias, llegando a cinco sapos de diferentes colores (rojo, naranja, amarillo, verde claro y verde oscuro), las que se observan en la Ilustración 1 (imagen de la derecha). Notar que las evaluaciones tenían además caras asociadas al sentimiento percibido en el viaje, por ejemplo, la roja era de enojo/malestar y la verde era feliz/contento.

Dado lo anterior, se procede a normalizar las evaluaciones a los valores 1, 2, 3, 4 y 5. En las evaluaciones previas el 2 y 4 están ausentes, a diferencia de las nuevas. En Anexos, Apéndice I: Notas de la base de datos, se muestra la tabla que indica qué datos fueron convertidos a cuál evaluación.

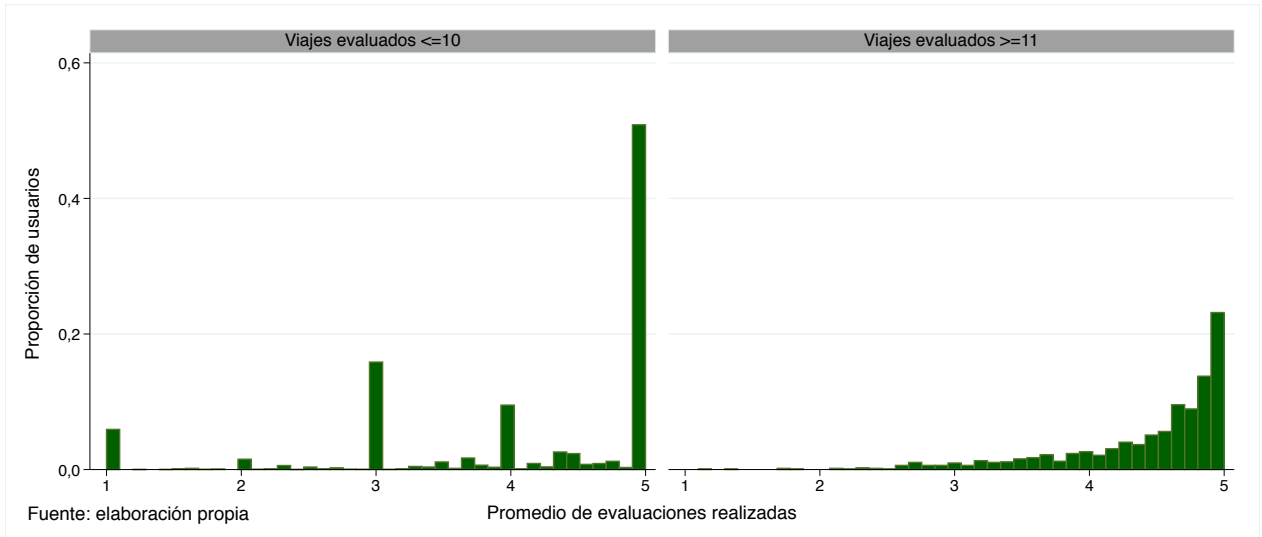
En los siguientes capítulos se trabaja como si la nota de la encuesta fuese continua, para así obtener variaciones y/o fluctuaciones, no obstante, en el subcapítulo: “Modelo para explicar variaciones en la nota del viaje”, se considera su naturaleza ordinal categórica y se trabaja en esa línea.

3.4.3.1 Distribución de usuarios según promedio de evaluación

De manera general, en la Ilustración 28 se grafican dos histogramas para observar la distribución de usuarios según la nota promedio de todos sus viajes evaluados. Se diferencia dos tipos de usuarios según la Tabla 23, debido a que en el gráfico de la izquierda se considera los usuarios que evaluaron menos o igual a 10 viajes y que representan el 94,06% de los usuarios evaluadores. Por el otro lado, en el gráfico de la derecha son quienes evaluaron más de 11 viajes, y representan cerca del 6% de los evaluadores.

¹⁷ El sapo (animal) es la mascota de la aplicación.

Ilustración 28: Histogramas de evaluación promedio, diferenciando usuarios según cantidad de viajes evaluados.

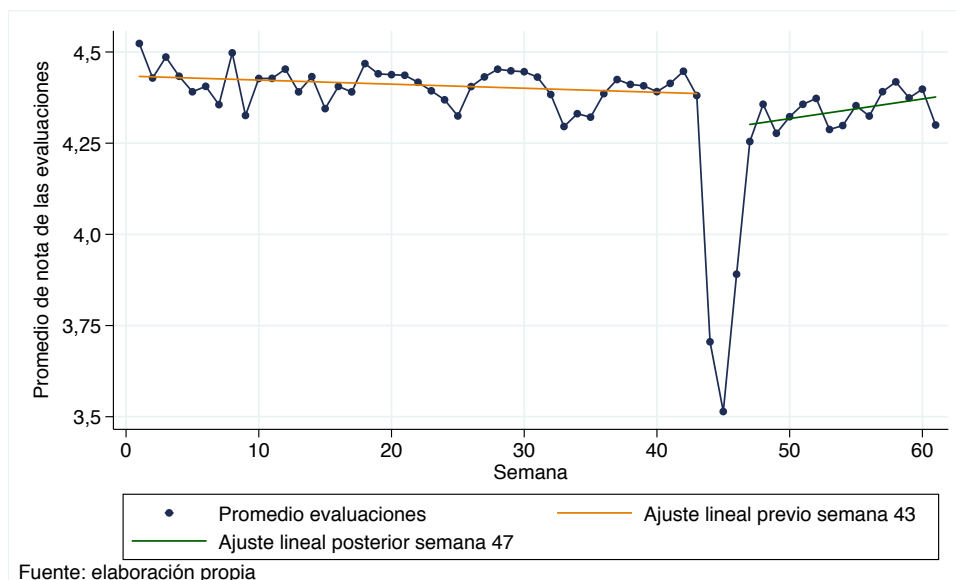


Se observa cómo la distribución de usuarios en general indica notas mayormente positivas de los viajes (cuya nota promedio es mayor o igual a 3), incluso en ambos tipos de usuarios la cantidad más importante de usuarios (en proporción) tienen nota promedio de 5.

3.4.3.2 Variación temporal de la evaluación

A continuación se grafica la evaluación promedio de los viajes evaluados según semana (ver Ilustración 29), para evidenciar variaciones relevantes en la nota.

Ilustración 29: Variación semanal de la evaluación promedio de los viajes evaluados.



Se observa que la evaluación promedio semanal de los viajes está mayoritariamente entre 4,25 y 4,5, valores positivos en términos de evaluación. La anomalía de menor evaluaciones realizadas

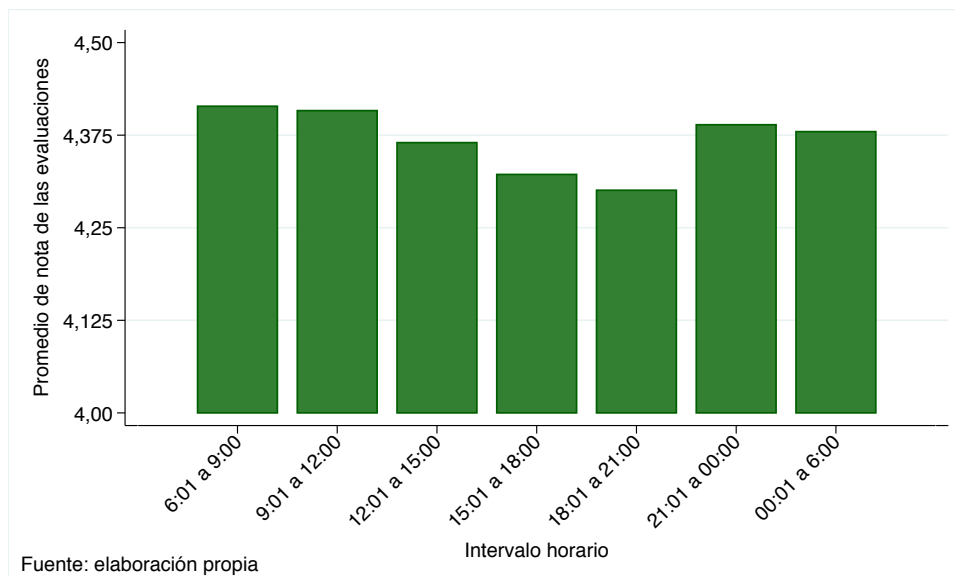
(observada en la Ilustración 22 e Ilustración 24) también está presente, provocando que la evaluación promedio particularmente durante este período sea más baja.

Al incluir un ajuste lineal a todo el gráfico, se representa una línea temporal con pendiente negativa más profunda que la naranja, debido al efecto de las semanas 44-46. Por esto es que se realizan dos ajustes lineales, uno previo y posterior. El ajuste lineal previo hasta la semana 43 en general es bastante estable a pesar de que tiene una tendencia ligeramente negativa. Por el otro lado, el ajuste lineal posterior a la semana 47 tiene una pendiente más fuerte positiva. En general, no es inmediato identificar un patrón de cómo varía la nota según la semana.

3.4.3.3 Variación temporal de la evaluación, según hora del día

Luego, se busca identificar si la evaluación de los viajes presenta variaciones según la hora del día. Para esto se grafica la evaluación promedio según los mismos intervalos horarios utilizados en los capítulos previos de reportes y consultas de paraderos (ver Ilustración 30).

Ilustración 30: Variación de la evaluación promedio de los viajes evaluados según intervalo horario.

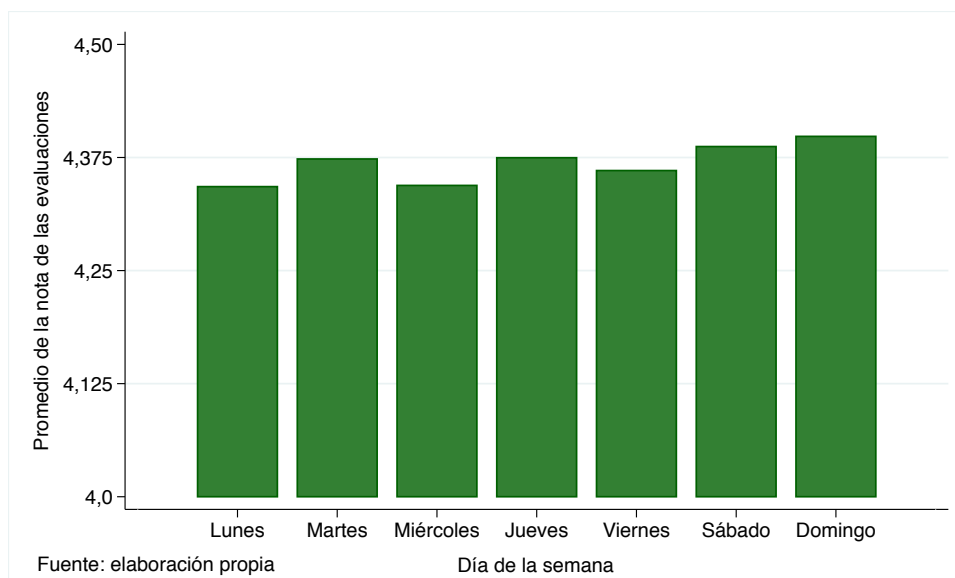


En la Ilustración 30 se observa una ligera variación de la nota promedio según el intervalo horario del día. Visualmente el mejor evaluado es la hora punta de la mañana y en contraste, el peor evaluado es el de la punta parte.

3.4.3.4 Variación temporal de la evaluación, según día de la semana

De igual forma, se busca analizar si la nota también varía según el día de la semana, para lo cual se grafica la evaluación promedio en la siguiente Ilustración 31.

Ilustración 31: Variación de la nota promedio de los viajes evaluados según día de la semana.



Se observa que el promedio de las evaluaciones realizadas según el día de la semana también presenta ligeras variaciones. Destaca que los días mejores evaluados son el sábado y domingo, en cambio, el peor evaluado es el lunes.

3.4.3.5 Variación local de la evaluación, según duración y tiempo del viaje

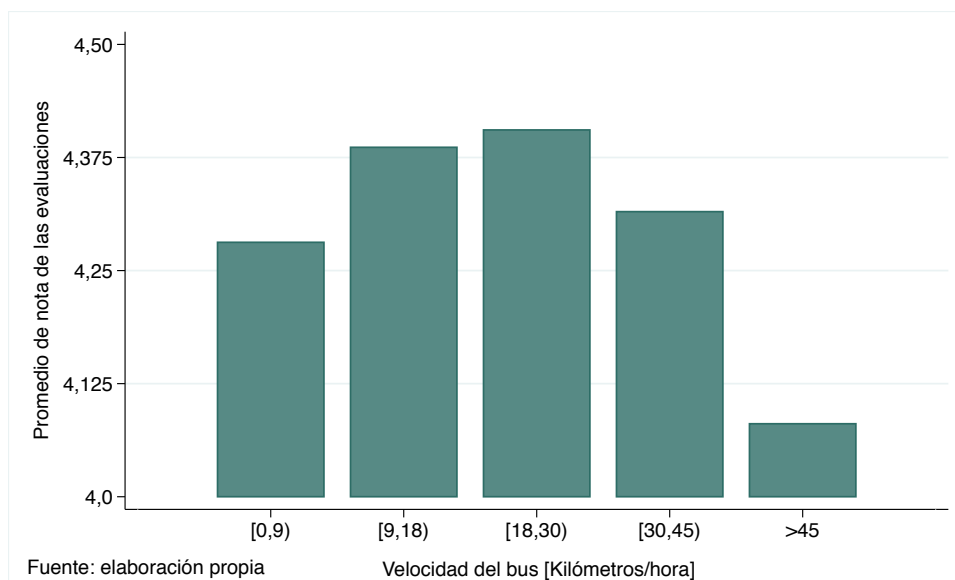
En este capítulo se ahonda en algo más técnico que es analizar si la nota del viaje está afectada por la duración y tiempo del viaje. Dado que ambas variables son importantes en el análisis, se utiliza la velocidad del bus calculada a partir de estas variables. En la Tabla 28 se muestra la cantidad de evaluaciones y el promedio de las notas según la velocidad del bus.

Tabla 28: Cantidad de evaluaciones y promedio de la nota según velocidad del viaje evaluado.

| Velocidad del recorrido [Mts/minuto] | Velocidad del recorrido [Km/hr] | Cantidad de evaluaciones | Promedio de evaluación |
|--------------------------------------|---------------------------------|--------------------------|------------------------|
| Inferior a 150 | Inferior a 9 | 13.337 | 4,2812 |
| Entre 150 y 300 | Entre 9 y 18 | 38.635 | 4,3862 |
| Entre 300 y 500 | Entre 18 y 30 | 11.456 | 4,4054 |
| Entre 500 y 750 | Entre 30 y 45 | 1.025 | 4,3151 |
| Superior a 750 | Superior a 45 | 372 | 4,0806 |

En la Ilustración 32 se utiliza esta información y se muestra las variaciones de la nota según la velocidad del bus. Lo esperable sería encontrar una curva creciente, donde aquellos buses más rápidos debiesen ser mejor evaluados, no obstante, aquellos superiores a 30 kilómetros por hora tienen una curva decreciente en su evaluación promedio.

Ilustración 32: Variación de la nota promedio de las evaluaciones según velocidad del bus.



3.4.4 Modelo para explicar variaciones en la nota del viaje

Dado que la variable dependiente, nota de la evaluación, es categórica y ordenada, se utiliza un modelo probit ordinal. No obstante, hay que aplicar dos modelos, uno cuando las posibles notas eran 1-3-5 y otro cuando eran 1-2-3-4-5, esto debido a que el modelo explica probabilidad y subestimaría las notas 2 y 4 al estar menos presentes en la muestra.

La primera evaluación diferente al patrón 1-3-5 es el 20 de mayo, por lo que se separa la regresión en antes y posterior a esa fecha. A continuación se presenta estadísticas básicas según período en la Tabla 29.

Tabla 29: Cantidad de evaluaciones y estadísticas según formato de evaluación.

| Fecha | Viajes evaluados | Porcentaje | Acumulado |
|-----------------------------------|------------------|----------------|-----------|
| 1-Agosto-2018 a 19-Mayo-2019 | 47.615 | 26,55% | 26,55% |
| 20-Mayo-2019 a 30-Septiembre-2019 | 17.210 | 73,45% | 100,00% |
| Total | 64.825 | 100,00% | - |

En ambos modelos, las variables independientes a estudiar son los días de la semana, el intervalo horario (mediante formato binario, donde es 1 si es el horario indicado y 0 si no) y si el usuario es viajero frecuente (en la siguiente parte se aborda este concepto). Se omite la variable velocidad ya que está correlacionada fuertemente con las variables anteriores (se realiza un análisis de cómo la velocidad, el cual está adjunto en Anexos, Apéndice J: Velocidad del viaje según día de la semana e intervalo horario, tiene diferencias estadísticamente significativas dependiendo de estos atributos) y al ser incluida en el modelo junto al tipo de usuario, no es significativa en ambos modelos.

3.4.4.1 Modelo evaluación de tres evaluaciones, primer período

En la Tabla 30 se adjunta la cantidad de evaluaciones existentes por cada evaluación ingresada, con los posibles valores 1, 2 y 3. Se tiene que la más frecuente es la evaluación 5, con cerca del 77% de los datos.

Tabla 30: Estadísticas de las evaluaciones de formato de 3 evaluaciones.

| Evaluación | Frecuencia | Porcentaje | Acumulado |
|--------------|---------------|----------------|-----------|
| 1 | 3.207 | 6,74% | 6,74% |
| 3 | 7.762 | 16,30% | 23,04% |
| 5 | 36.646 | 76,96% | 100,00% |
| Total | 47.615 | 100,00% | - |

Se busca diferenciar a los usuarios que viajan más que el resto de los usuarios, por lo que se denomina “Viajeros frecuentes”. A partir de la Tabla 31, se considerará un viajero frecuente a quien tenga más de 10 viajes en el primer período, que equivalen al 8,9% de los usuarios viajeros.

Tabla 31: Distribución de usuarios según viajes realizados en período 1.

| Viajes realizados | Frecuencia | Porcentaje | Acumulado |
|-------------------|---------------|----------------|-----------|
| Sólo 1 | 5.337 | 49,72% | 49,72% |
| Entre 2 y 9 | 4.445 | 41,4% | 91,13% |
| Superior a 10 | 953 | 8,9% | 100,00% |
| Total | 10.735 | 100,00% | - |

Dado lo anterior, se estima un probit ordinal, cuya variable dependiente es la evaluación recibida y las variables independientes son si el usuario es viajero frecuente, los días de la semana y el intervalo horario. El grupo base para las últimas dos variables serán el día lunes y 18:01 a 21:00 dado que son los momentos en los que la evaluación es más baja como se vió anteriormente. Los resultados de este modelo se encuentran en la Tabla 32.

Tabla 32: Resultados estimación modelo probit ordenado de evaluación para 3 evaluaciones.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| Viajero frecuente | 0,3600 | 0,012 | 28,92 | 0,000 |
| Martes | 0,0424 | 0,021 | 1,93 | 0,053 |
| Miércoles | -0,0227 | 0,021 | -1,06 | 0,288 |
| Jueves | 0,0348 | 0,021 | 1,59 | 0,111 |
| Viernes | 0,0022 | 0,021 | 0,10 | 0,920 |
| Sábado | 0,0609 | 0,024 | 2,45 | 0,014 |
| Domingo | 0,0778 | 0,027 | 2,85 | 0,004 |
| 6:01 a 9:00 | 0,1479 | 0,019 | 7,59 | 0,000 |
| 9:01 a 12:00 | 0,1262 | 0,021 | 5,83 | 0,000 |
| 12:01 a 15:00 | 0,0891 | 0,020 | 4,25 | 0,000 |
| 15:01 a 18:00 | 0,0545 | 0,020 | 2,67 | 0,008 |
| 21:01 a 00:00 | 0,0681 | 0,023 | 2,93 | 0,003 |
| 00:01 a 6:00 | 0,0770 | 0,045 | 1,70 | 0,088 |
| Intercepto 1 | -1,2296 | 0,021 | - | - |
| Intercepto 2 | -0,4584 | 0,020 | - | - |

Al interpretar los signos, se tiene que ser viajero frecuente (=1) es más probable estar en las categorías asociadas a mejores evaluaciones. Respecto a los días (=1 según corresponda) martes, sábado y domingo es más probable que las evaluaciones estén en mejores categorías respecto al día lunes (el resto de los días no tiene diferencia significativa). Finalmente, según el intervalo horario (=1 según corresponda) todos son significativos al 1%, excepto la madrugada (al 10%) y se tiene que es más probable que estén asociadas a categorías de evaluaciones más positivas.

Luego, en la Tabla 33 se adjunta el porcentaje de evaluaciones (que es el indicado en la Tabla 30) y la probabilidad predicha por el modelo probit ordinal.

Tabla 33: Probabilidades predichas por el modelo de 3 evaluaciones y el de los datos.

| Evaluación | 1 | 3 | 5 |
|-----------------------------------|----------|----------|----------|
| Porcentaje de evaluaciones | 6,74% | 16,30% | 76,96% |
| Probabilidad del modelo | 6,73% | 16,28% | 76,97% |

Para facilitar el entendimiento de los resultados del modelo, se obtiene los efectos marginales que se adjuntan en la Tabla 34. El cambio en las variables binarias es medido de 0 a 1.

Tabla 34: Resultados efectos marginales de probit ordenado de 3 evaluaciones.

| Coeficiente | Pr(Evaluación=1) | | Pr(Evaluación=3) | | Pr(Evaluación=5) | |
|-------------------|------------------|----------|------------------|----------|------------------|----------|
| | Efecto marginal | Pr(> z) | Efecto marginal | Pr(> z) | Efecto marginal | Pr(> z) |
| Viajero frecuente | -4,60% | 0,000 | -6,27% | 0,000 | 10,87% | 0,000 |
| Martes | -0,51% | 0,048 | -0,74% | 0,052 | 1,26% | 0,050 |
| Miércoles | 0,28% | 0,294 | 0,39% | 0,289 | -0,68% | 0,291 |
| Jueves | -0,42% | 0,105 | -0,60% | 0,110 | 1,03% | 0,108 |
| Viernes | -0,02% | 0,920 | 0,03% | 0,920 | 0,06% | 0,920 |
| Sábado | -0,72% | 0,011 | -1,04% | 0,014 | 1,77% | 0,013 |
| Domingo | -0,92% | 0,003 | -1,35% | 0,004 | 2,28% | 0,003 |
| 6:01 a 9:00 | -1,73% | 0,000 | -2,56% | 0,000 | 4,30% | 0,000 |
| 9:01 a 12:00 | -1,47% | 0,000 | -2,19% | 0,000 | 3,66% | 0,000 |
| 12:01 a 15:00 | -1,06% | 0,000 | -1,55% | 0,000 | 2,61% | 0,000 |
| 15:01 a 18:00 | -0,66% | 0,006 | -0,95% | 0,007 | 1,61% | 0,007 |
| 21:01 a 00:00 | -0,81% | 0,002 | -1,19% | 0,003 | 2,00% | 0,003 |
| 00:01 a 6:00 | -0,91% | 0,071 | -1,34% | 0,085 | 2,25% | 0,079 |

Se concluye que ser un viajero frecuente aumenta significativamente en un 10,87% la probabilidad de evaluar 5, y disminuye significativamente en 4,60% y 6,27% la probabilidad de evaluar 1 y 3, respectivamente, respecto al grupo base (quienes evalúan poco frecuentemente). Se concluye así que los viajeros frecuentes tienden a evaluar más positivamente sus viajes, lo cual podría deberse a un potencial sesgo de selección.

Respecto a los días, sólo los coeficientes de los días Martes, Sábados y Domingos presentan diferencias significativas respecto al día Lunes. Se tiene que es menos probable evaluar en las categorías 1 y 3 éstos días, y más probable en la categoría 5, respecto al día Lunes. Sobre el intervalo horario, todas las diferencias son significativas respecto a la hora punta de la tarde (notar que la madrugada es significativa al 10%). Se desprende que en todos los horarios es menos probable evaluar en las categorías 1 y 3, y más probable evaluar en la categoría 5.

A modo general se tiene que los días efectivamente afectan la evaluación del viaje, al igual que el intervalo horario. La magnitud de la diferencia varía entre -2% y 3% en el día del viaje, y entre -2% y 5% en el intervalo horario, los cuales son más bajos al observado del ser un viajero frecuente, que varía entre -7% y 11%.

3.4.4.2 Modelo evaluación de cinco evaluaciones, segundo período

En la Tabla 35 se adjunta la cantidad de evaluaciones existentes por cada evaluación ingresada, con los posibles valores de 1 a 5. Se tiene que la más frecuente es la categoría 5, con cerca del 57% de las evaluaciones.

Tabla 35: Estadísticas de las evaluaciones de formato de 5 evaluaciones.

| Evaluación | Frecuencia | Porcentaje | Acumulado |
|-------------------|-------------------|-------------------|------------------|
| 1 | 821 | 4,77% | 4,77% |
| 2 | 454 | 2,64% | 7,41% |
| 3 | 2.065 | 12,00% | 19,41% |
| 4 | 4.025 | 23,39% | 42,79% |
| 5 | 9.845 | 57,21% | 100,00% |
| Total | 17.210 | 100,00% | - |

De igual forma que en el subcapítulo previo 3.4.4.1, se diferencia el tipo de usuario “Viajeros frecuente”. Para esto, se considerará un usuario frecuente a quien tenga más de 10 viajes en el segundo período, que equivalen al 7,4% de los usuarios viajeros. La frecuencia se muestra en la Tabla 36.

Tabla 36: Distribución de usuarios según viajes realizados en período 1.

| Viajes realizados | Frecuencia | Porcentaje | Acumulado |
|--------------------------|-------------------|-------------------|------------------|
| Sólo 1 | 2.244 | 51,41% | 51,41% |
| Entre 2 y 9 | 1.797 | 41,17% | 92,58% |
| Superior a 10 | 324 | 7,42% | 100,00% |
| Total | 4.365 | 100,00% | - |

Dado lo anterior, se estima un probit ordinal, cuya variable dependiente es la evaluación recibida y las variables independientes son si el usuario es viajero frecuente, los días de la semana y el intervalo horario. El grupo base para las últimas dos variables serán el día lunes y 18:01 a 21:00 dado que son los momentos en los que la evaluación es más baja como se vió anteriormente. Los resultados se adjuntan en la Tabla 37.

Tabla 37: Resultados estimación modelo probit ordenado de evaluación para 5 evaluaciones.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| Viajero frecuente | 0,3995 | 0,017 | 22,26 | 0,000 |
| Martes | 0,0086 | 0,030 | 0,28 | 0,777 |
| Miércoles | 0,0386 | 0,030 | 1,26 | 0,207 |
| Jueves | 0,0505 | 0,030 | 1,65 | 0,099 |
| Viernes | 0,0682 | 0,030 | 2,21 | 0,027 |
| Sábado | 0,1132 | 0,034 | 3,27 | 0,001 |
| Domingo | 0,1372 | 0,038 | 3,56 | 0,000 |
| 6:01 a 9:00 | -0,0434 | 0,027 | -1,57 | 0,117 |
| 9:01 a 12:00 | 0,0623 | 0,031 | 2,00 | 0,046 |
| 12:01 a 15:00 | 0,0224 | 0,030 | 0,74 | 0,457 |
| 15:01 a 18:00 | -0,0257 | 0,029 | -0,89 | 0,375 |
| 21:01 a 00:00 | 0,0962 | 0,034 | 2,80 | 0,005 |
| 00:01 a 6:00 | 0,1095 | 0,063 | 1,71 | 0,087 |
| Intercepto 1 | -1,2296 | 0,031 | - | - |
| Intercepto 2 | -0,4584 | 0,030 | - | - |
| Intercepto 3 | -0,6243 | 0,029 | - | - |
| Intercepto 4 | 0,0721 | 0,029 | - | - |

Al interpretar los signos, se tiene que ser viajero frecuente (=1) es más probable estar en las categorías asociadas a mejores evaluaciones. Respecto a los días (=1 según corresponda) sólo los Viernes, Sábados y Domingo son significativos al 1% y al tener coeficientes positivos significa que es más probable que estén asociado a categorías de mejores evaluaciones respecto al lunes, grupo base. De igual forma, según el intervalo horario (=1 según corresponda) sólo es significativo el intervalo 9:01 a 12:00 y 21:01 a 00:00, y tienen coeficientes positivos lo que implica que es más probable que estén asociados a mejores evaluaciones, respecto al horario punta de la tarde.

Luego, en la Tabla 38 se adjunta el porcentaje de evaluaciones (que es el indicado en la Tabla 35) y la probabilidad predicha por el modelo probit ordenado anterior, las cuales coinciden de buena forma.

Tabla 38: Probabilidades predichas por el modelo de 5 evaluaciones y el de los datos.

| Evaluación | 1 | 2 | 3 | 4 | 5 |
|-----------------------------------|----------|----------|----------|----------|----------|
| Porcentaje de evaluaciones | 4,77% | 2,64% | 12,00% | 23,39% | 57,21% |
| Probabilidad del modelo | 4,77% | 2,63% | 11,99% | 23,39% | 57,19% |

En la Tabla 39 se adjuntan los efectos marginales, en donde la variación en las variables discretas es de 0 a 1.

Tabla 39: Resultados efectos marginales de probit ordenado de 5 evaluaciones.

| Coeficiente | Pr(Evaluación=1) | | Pr(Evaluación=2) | | Pr(Evaluación=3) | | Pr(Evaluación=4) | | Pr(Evaluación=5) | |
|-----------------|------------------|----------|------------------|----------|------------------|----------|------------------|----------|------------------|----------|
| | Efecto marginal | Pr(> z) | Efecto marginal | Pr(> z) | Efecto marginal | Pr(> z) | Efecto marginal | Pr(> z) | Efecto marginal | Pr(> z) |
| Viaj. frecuente | -3,76% | 0,000 | -1,61% | 0,000 | -5,39% | 0,000 | -4,80% | 0,000 | 15,56% | 0,000 |
| Martes | -0,08% | 0,775 | -0,03% | 0,776 | -0,12% | 0,776 | -0,11% | 0,777 | 0,34% | 0,776 |
| Miércoles | -0,35% | 0,197 | -0,15% | 0,202 | -0,53% | 0,205 | -0,48% | 0,215 | 1,51% | 0,206 |
| Jueves | -0,45% | 0,089 | -0,20% | 0,094 | -0,69% | 0,097 | -0,63% | 0,106 | 1,97% | 0,097 |
| Viernes | -0,61% | 0,021 | -0,27% | 0,024 | -0,92% | 0,026 | -0,86% | 0,032 | 2,66% | 0,026 |
| Sábado | -0,98% | 0,000 | -0,44% | 0,001 | -1,52% | 0,001 | -1,46% | 0,002 | 4,40% | 0,001 |
| Domingo | -1,16% | 0,000 | -0,52% | 0,000 | -1,83% | 0,000 | -1,79% | 0,001 | 5,31% | 0,000 |
| 6:01 a 9:00 | 0,41% | 0,126 | 0,18% | 0,123 | 0,59% | 0,119 | 0,52% | 0,111 | -1,71% | 0,118 |
| 9:01 a 12:00 | -0,56% | 0,038 | -0,25% | 0,042 | -0,84% | 0,044 | -0,78% | 0,052 | 2,43% | 0,045 |
| 12:01 a 15:00 | -0,20% | 0,451 | -0,09% | 0,454 | -0,31% | 0,456 | -0,28% | 0,461 | 0,88% | 0,456 |
| 15:01 a 18:00 | 0,24% | 0,382 | 0,10% | 0,380 | 0,35% | 0,377 | 0,31% | 0,370 | -1,01% | 0,376 |
| 21:01 a 00:00 | -0,84% | 0,003 | -0,38% | 0,004 | -1,30% | 0,005 | -1,23% | 0,007 | 3,74% | 0,005 |
| 00:01 a 6:00 | -0,93% | 0,060 | -0,42% | 0,071 | -1,47% | 0,080 | -1,42% | 0,106 | 4,25% | 0,083 |

Se concluye que ser viajero frecuente aumenta significativamente en un 15,56% la probabilidad de evaluar con nota 5, y disminuye significativamente la probabilidad de evaluar con nota 1-4 (rango de -5,39% a -1,61%). Al igual que en el modelo anterior, los viajeros frecuentes tienden a evaluar más positivamente sus viajes.

Respecto a los días, sólo los coeficientes de los días Martes, Sábados y Domingos indican diferencias significativas respecto al día Lunes. En donde es menos probable evaluar en las categorías 1, 2, 3 y 4 éstos días, y es más probable en la categoría 5, todo respecto al día Lunes.

Sobre el intervalo horario, sólo el intervalo de 9:01 a 12:00 y 21:01 a 00:00 son significativos. Se observa así que en estos horarios es menos probable evaluar en las categorías 1 a 4, pero es más probable evaluar en la categoría 5, respecto a la hora punta de la tarde.

La conclusión es similar al primer período, los días y el intervalo horario también afectan la nueva evaluación del viaje.

3.4.4.3 Conclusiones generales del modelo

Notar que si bien existen períodos horarios y días de la semana que tienen una diferencia significativa en la probabilidad de ser evaluado en las diferentes categorías, esto no implica que aquella diferencia sea debido a que el viaje sea en ése período/día, sino que se debe a no observables que están correlacionados y son variables omitidas. Por ejemplo, la temperatura del bus (depende de la hora y mes del año), la congestión vehicular (depende del día y la hora), la densidad de pasajeros (depende del día y hora también), entre otras.

Lo que sí es concluyente e importante, es que aquellos usuarios que suelen evaluar frecuentemente sus viajes, son más propensos a evaluar positivamente sus viajes (que aquellos que no lo son), lo cual también puede dar indicio de una simultaneidad en donde quizás son usuarios evaluadores frecuentes porque son más sensibles a la la calidad del servicio.

3.5 Conclusiones del trabajo de datos

Si bien el trabajo de este capítulo ha sido mayoritariamente descriptivo, permite identificar aspectos generales de las acciones de los usuarios a lo largo del período estudiado. En los siguientes párrafos se enuncia los principales aprendizajes según base de datos analizada.

Respecto del trabajo con los datos de las consultas únicas de paraderos, se destaca la semanas de actividad de los usuarios, en donde se identificó que sólo el 34,9% de los usuarios que alguna vez consulta lo sigue haciendo posterior a un mes, y esta cifra se reduce a un 14,6% después de tres meses. Se desconoce si estas cifras son competitivas respecto a la competencia (para así concluir si son positivas o preocupantes). Si bien la cantidad de usuarios activos crece de manera estable en el período, existe un constante recambio de usuarios dada las cifras de permanencia en la aplicación mencionadas previamente. Por otro lado, el 80% de los usuarios consultores realiza menos de 250 consultas únicas y el 3% realiza sobre 2.500, lo que evidencia un efecto heterogéneo en los usuarios, al existir una gran proporción que presenta una baja actividad y una baja proporción que presenta alta actividad. Acerca de las consultas propiamente tal, se tiene que la cantidad promedio por usuario ha disminuido en el período estudiado, y que, esperablemente, se realizan más en los horarios punta y en los días laborales.

De la base de datos de reportes, primero se tiene la actividad de los usuarios, en donde el 93% de ellos tiene menos de 25 reportes realizados, y que el promedio de reportes por usuario ha disminuido en el tiempo (al igual que las consultas únicas). Por otro lado, respecto a la categoría de reportes, en primer lugar, la proporción de reportes de buses y paraderos era similar y estable hasta Marzo de 2019, cuando la proporción de paraderos se vuelve cada vez más dominante; en segundo lugar, la proporción de reportes positivos y pesimistas es estable en el tiempo y fluctúa entre 40/50% y 50/60% respectivamente; tercero y último, si bien la cantidad de reportes afirmativos es dominante sobre los negativos, la proporción de éstos del total se reduce progresivamente. Al estudiar las características locales, se tiene que los usuarios emiten más reportes en las horas punta (que además es cuando el promedio de reportes por usuario es el más alto) y en los días laborales, lo cual también es esperable. Finalmente, se destaca y comprueba el problema de participación en el cual *Transapp* también es caso de estudio de la Regla de Nielsen de 90:9:1, en donde al considerar todo el período de estudio y las estadísticas mensuales, en promedio, el 83% de los usuarios no reporta, el 15% de los usuarios reporta el 52% de los reportes, y el 1% restante realiza el 48% de los reportes remanentes.

En relación a la última base de datos estudiada, la de viajes y evaluaciones realizadas, se indica que si bien la cantidad promedio de viajes por usuario se mantiene estable en el tiempo, la cantidad promedio de evaluaciones por usuario aumenta, lo cual suena positivo, pero al considerar la situación completa no lo es, ya que la proporción de usuarios evaluadores disminuye en el tiempo, y quienes siguen evaluando, evalúan cada vez más. Respecto a las estadísticas de usuarios, se mantiene los patrones de las base de datos previas: el 80% de los usuarios viajeros ha viajado entre 1 y 4 veces, cifra que sube a 85% en los evaluadores. De igual manera, y esperablemente, en las

horas punta es donde se realizan más viajes y evaluaciones en el día. Destaca que la proporción de viajes evaluados es significativamente más alta (comprobado estadísticamente) durante la mañana y posterior a la hora punta de la tarde. Un aspecto inesperado es que la nota promedio según velocidad del bus no sigue un patrón monótono creciente, si no que es una curva creciente que llega a un peak y disminuye posterior a los 30 kilómetros por hora. Finalmente, se realizó un modelo para identificar qué atributos explicaban variaciones en la evaluación del viaje asignada por los usuarios, en donde se determinó que varía según la hora del día y el día de la semana, no obstante, esto se debe a que éstas variables absorben efectos de variables no observables como la congestión, el hacinamiento, la velocidad, temperatura y/o la seguridad (Eboli & Mazzullaba, 2007) (De Oña, De Oña, Eboli, & Mazzullaba, 2013). Lo que sí fue concluyente y lo importante del estudio, es que aquellos usuarios que evalúan frecuentemente viajes en la aplicación, son más propensos a realizar evaluaciones más positivas y menos negativas de sus viajes; evidenciando así un potencial sesgo de selección de estos usuarios.

Capítulo 4

Análisis de sobrevivencia de usuarios

El concepto “valor del tiempo de vida del usuario” (*Customer lifetime value*, CLV) en marketing es definido por Pfeifer, Haskins & Conroy (2005) como el valor presente de los flujos monetarios futuros de la relación con los consumidores. Tiene tres principales componentes en su estimación: primero, el valor del consumidor a través del tiempo; segundo, la tasa de descuento (del dinero); y tercero, la retención (probabilidad) de los usuarios en el tiempo, la cual puede modelarse utilizando una función de sobrevivencia (Rosset, Neumann, Eick, Vatnik, & Idan, 2002).

Es importante luego tener claridad de la retención de usuarios en la aplicación, y como se ve en el capítulo anterior, es del 45% de los usuarios posterior a los 4 meses, por lo que en este capítulo se busca profundizar utilizando un análisis de sobrevivencia, en donde el evento de falla corresponde a la semana de la última consulta de paradero del usuario (inicio de la inactividad), y el tiempo de sobrevivencia se define como la cantidad de semanas entre la primera y última consulta del usuario (período de actividad).

En este análisis, el tiempo de sobrevivencia del usuario es estático y se mide desde el inicio de la actividad de cada usuario, sin ser afectado por el transcurrir del tiempo. Por ejemplo, si el evento de falla (dejar de consultar) ocurre en la cuarta semana a dos individuos, pero uno es en abril y otro en octubre, ambos tienen un tiempo de sobrevivencia de 4 semanas.

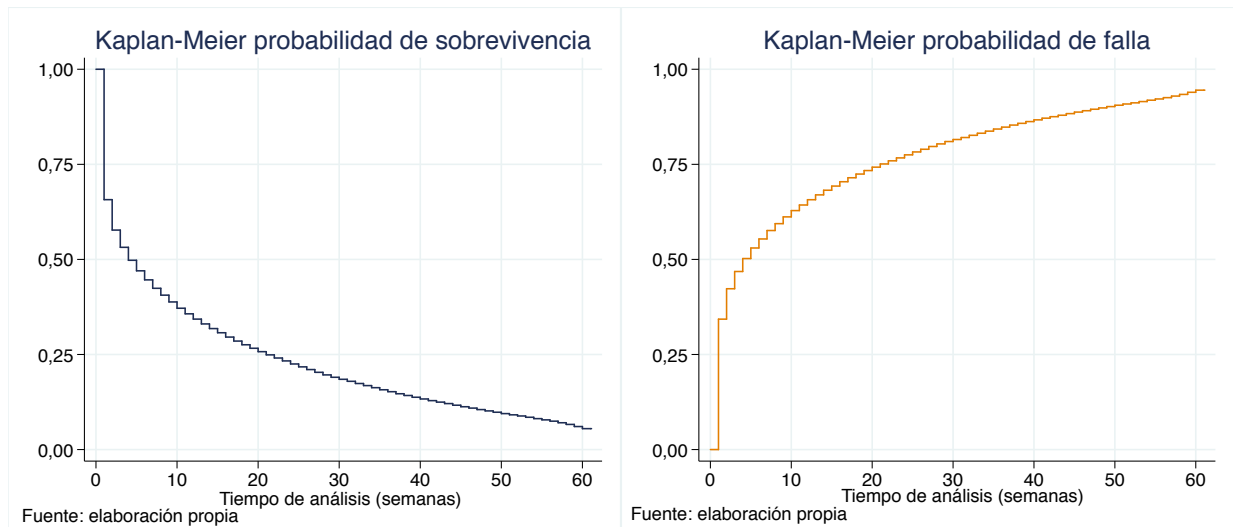
Al considerar los 152.312 usuarios, se tiene una tasa de incidencia de 0,767, es decir, el 76,7% de los usuarios deja de ser activo previo a la semana 61 (última del estudio), y el restante, se considera usuarios censurados, es decir, que hasta el término de la fecha del estudio el evento de falla (dejar consultar) aún no ocurre. Notar que a todos los usuarios se les conoce la fecha final de su última consulta, y que sólo a quienes dejaron de consultar en la 61 se les considera censurados, por lo que hay que considerar que probablemente aquellos usuarios que consultaron en la semana 59/60 (y que no consultaron en la 61), también potencialmente sigan siendo usuarios activos.

Tabla 40: Resumen estadísticas de modelación tiempo de sobrevivencia.

| Tiempo en riesgo [semanas] | Tasa de incidencia | Número de individuos | Tiempo de sobrevivencia | | |
|----------------------------|--------------------|----------------------|-------------------------|-----|-----|
| | | | 25% | 50% | 75% |
| 1.565.915 | 0,767 | 152.312 | 1 | 4 | 21 |

En la siguiente página se adjunta los gráficos Kaplan-Meier de probabilidad de sobrevivencia y de falla de los usuarios según la cantidad de semanas activos (notar que son inversos, ya que la probabilidad de falla se define como 1-probabilidad de sobrevivencia) (Ilustración 33).

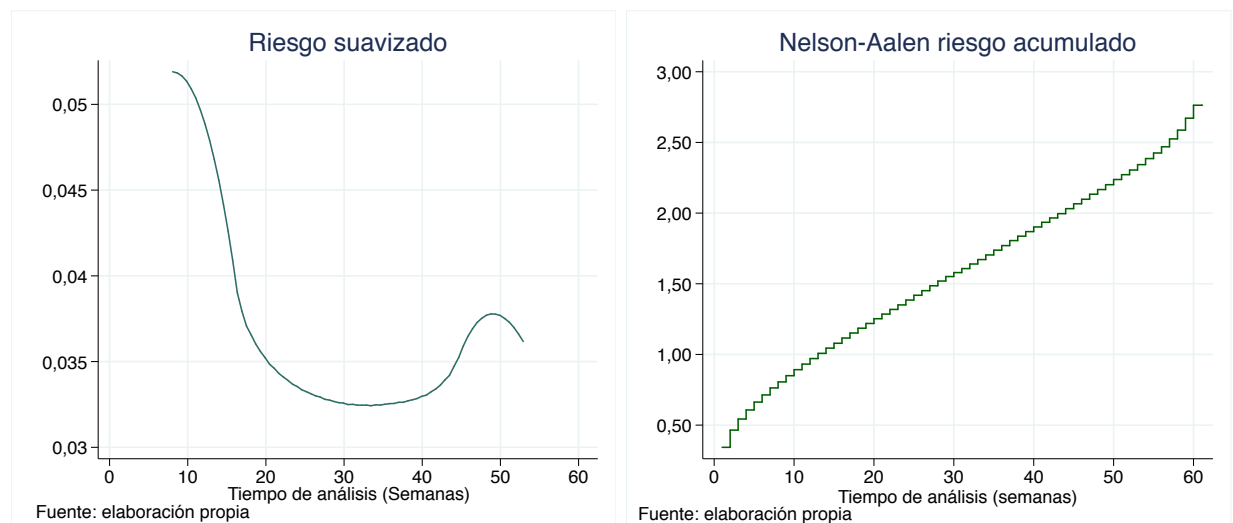
Ilustración 33: Probabilidades de supervivencia (izquierda) y de falla (derecha) de Kaplan-Meier considerando todos los usuarios que consultan tiempos de espera en paraderos en el período.



Del gráfico de la derecha, probabilidad de falla, se observa los puntos percentiles de la tabla anterior, en donde el usuario del percentil 25 es activo hasta la semana 1, el 50 hasta la semana 4 y el 75 hasta la semana 21.

En la gráfica de la izquierda, se tiene el riesgo suavizado, que se interpreta como la probabilidad de falla instantánea de la semana condicionado a el evento de falla aún no ocurre (Cleves, Gould, & Marchenko, 2008), en donde se observa una parte decreciente y creciente. Se desprende que una vez que los usuarios superan una cierta cantidad de semanas (por ejemplo, 20), su probabilidad de dejar de consultar disminuye, lo cual es esperable ya que simbolizaría que generan una relación de confianza y hábito con la aplicación en sus viajes en la ciudad, a diferencia de aquellos que sólo están un par de semanas.

Ilustración 34: Riesgo suavizado (izquierda) y riesgo acumulado de Nelson-Aalen (derecha) de los usuarios consultores según semanas de actividad.



Por el otro lado, en el gráfico de la derecha de la Ilustración 34, se tiene el riesgo acumulado (*cummulative hazard*) que es la suma (de hecho, la integral) del riesgo acumulado hasta el tiempo t dado que aún no ocurre el evento de falla (Cleves, Gould, & Marchenko, 2008). Se observa cómo el riesgo acumulado adquiere la información del riesgo suavizado (gráfico de la izquierda), ya que si bien crece constantemente el riesgo, se observa cómo este se desacelera progresivamente en las semanas iniciales y cómo se acelera en las semanas finales.

4.1 Análisis según actividades de la semana 1

¿Qué conlleva que un usuario que consulta un paradero siga utilizando la aplicación? Aquellos usuarios que utilizan más herramientas en la aplicación, es decir, reportan y viajan en su primera semana, ¿tienen menor riesgo de dejar de ser activos?. Para responder estas preguntas, primero se clasifican los usuarios según sus actividades realizadas en su primera semana de actividad¹⁸ (ver Tabla 41), en donde se genera tres variables binarias: Reportador (si reporta), Viajador (si viaja) y Ambos (si reporta y viaja). Donde todos los usuarios consultaron un tiempo de llegada de un bus.

Tabla 41: Clasificación de usuarios según acciones durante la primera semana de actividad de consulta de paraderos.

| Categoría | Cantidad | Porcentaje | Acumulado |
|-----------------------------|----------------|----------------|-----------|
| Reportador | 14.319 | 9,40% | 9,40% |
| Viajador | 3.243 | 2,13% | 11,53% |
| Ambos | 5.921 | 3,89% | 15,42% |
| Sólo consulta ¹⁹ | 128.829 | 84,58% | 100,00% |
| Total | 152.312 | 100,00% | - |

Se tiene así que el 84,58% de los usuarios en su primera semana de actividad (de consultas de paradero) sólo consultaron tiempos de espera, el 9,40% también reportó, el 2,13% además viajó, y el 3,89% realizó las tres acciones.

En la Tabla 42 se adjunta las estadísticas del modelo de análisis de sobrevivencia, en donde lo más importante se observa en la semana de evento de falla (inactividad) dada la categoría de usuario, en donde el percentil 50 está en la semana 5 para los que reportan en la primera semana, 4 para quienes realizan viajes en la primera semana, 6 para quienes realizan ambas actividades, y 4 para quienes sólo consultan.

¹⁸ Los usuarios que no están presentes en la base de datos de Consultas (es decir, no consultaron) y sólo reportaron (1.066) y/o viajaron (1.421) en su tiempo en la aplicación, son excluidos de este análisis, quienes totalizan 2.487 usuarios.

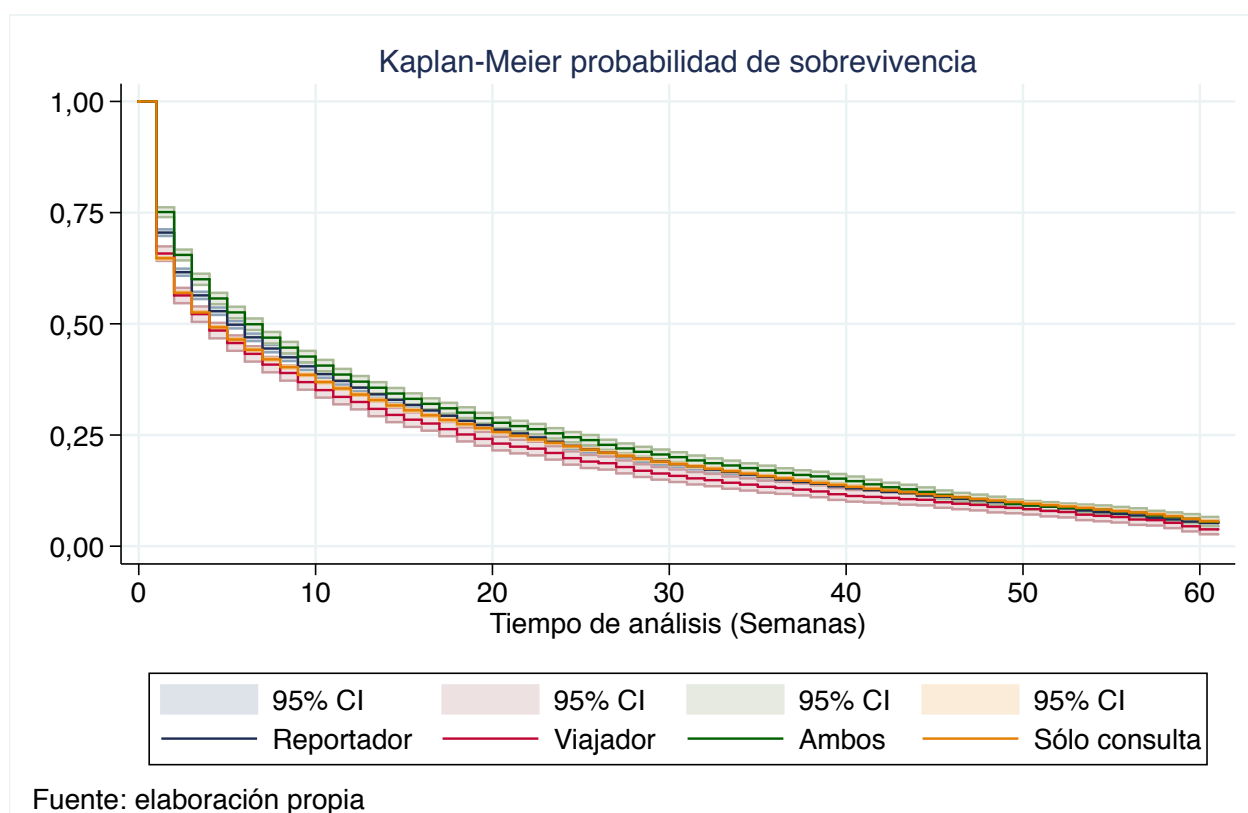
¹⁹ Los usuarios que reportaron (413) y/o viajaron (611) previo a la semana de su primera consulta, sí se consideran en el análisis, pero están agrupados en la categoría de que no viajan ni reportan en su primera semana. Totalizan 1.024 usuarios.

Tabla 42: Resumen estadísticas de modelación tiempo de sobrevivencia, según categoría de usuario en base a sus acciones durante la primera semana de actividad de consulta de paraderos.

| Categoría | Tiempo en riesgo | Tasa de incidencia | Tiempo de sobrevivencia | | |
|---------------|------------------|--------------------|-------------------------|-----|-----|
| | | | 25% | 50% | 75% |
| Reportador | 150.283 | 0,0744 | 1 | 5 | 22 |
| Viajador | 32.752 | 0,0818 | 1 | 4 | 19 |
| Ambos | 73.904 | 0,0659 | 2 | 6 | 24 |
| Sólo consulta | 1.308.976 | 0,0774 | 1 | 4 | 21 |

Luego la actividades realizadas en la primera semana pareciera indicar que aquellos usuarios tienen diferentes funciones de sobrevivencia y de riesgo, por lo que se procede a graficar la probabilidad de sobrevivencia de Kaplan-Meier junto con su intervalo de confianza al 95% (ver Ilustración 35).

Ilustración 35: Probabilidad de sobrevivencia de Kaplan-Meier según categoría de usuarios en base a actividades realizadas en primera semana de consulta de paraderos.



Se observa cómo ya en la segunda semana hay efectos diferenciados según las acciones realizadas en la primera semana. Se tiene que la tasa de sobrevivencia de quienes reportan y viajan (grupo Ambos, línea verde) es más alta que quienes sólo reportan (línea azul), y a su vez, de quienes viajan (línea roja) y sólo consultan (línea naranja), a pesar de que en estas últimas dos, no hay evidencia de diferencias entre ellas (debido a que los intervalos de confianza se cruzan).

En la Tabla 43 se adjunta los resultados del test log-rank para la igualdad de las funciones de supervivencia según la categoría de usuario. La hipótesis nula es que aquellas funciones son iguales.

Tabla 43: Test log-rank de la igualdad de las funciones de supervivencia según categoría de usuarios en base a acciones realizadas en la primera semana de consulta de paraderos.

| Categoría | Eventos observados | Eventos esperados | Test | |
|---------------|--------------------|-------------------|----------------|-------|
| Reportador | 11.186 | 11.523,72 | $\chi^2(3)$ | 57,94 |
| Viajador | 2.682 | 2.541,88 | $\Pr(>\chi^2)$ | 0,000 |
| Ambos | 4.874 | 5.249,89 | | |
| Sólo consulta | 101.383 | 100.809,51 | | |
| Total | 120.125 | 120.125,00 | | |

Dado el test de la tabla se rechaza la hipótesis nula de igualdad de las funciones de supervivencia, por lo que se concluye que las acciones realizadas por los usuarios en la primera semana de actividad de consultas sobre el tiempo de espera de buses, son determinante en la probabilidad de supervivencia como usuario activo en los siguientes períodos.

Al realizar una regresión Cox, que estima los efectos proporcionales del ratio de riesgo, con variable independientes los grupos de usuarios anteriores y grupo base quienes sólo consultan (grupo 4), se tienen los siguientes resultados (ver Tabla 44).

Tabla 44: Resultados regresión Cox.

| Categoría | Ratio de riesgo | Error estándar | Valor z | $\Pr(> z)$ |
|------------|-----------------|----------------|---------|-------------|
| Reportador | 0,9651 | 0,0096 | -3,56 | 0,000 |
| Viajador | 1,0491 | 0,0205 | 2,45 | 0,014 |
| Ambos | 0,9231 | 0,0135 | -5,46 | 0,000 |

Todas las covariables son significativas. De los coeficientes se desprende que el ratio de riesgo disminuye un 3,49% en los reportadores, aumenta en 4,91% en los viajeros y disminuye 7,69% en los reportadores y viajeros. Provocando así que la tasa de supervivencia sea más alta en aquellos donde disminuye el ratio de riesgo y más alta en aquellos donde aumenta, respecto al grupo base que son los usuarios que sólo consultan (consistente con lo observado en la Ilustración 35).

4.2 Análisis según actividades en el período

En este apartado se realiza un análisis similar, pero ahora en vez de concentrar el estudio en la primera semana de actividad, se clasifica a los usuarios (utilizando la misma categorización) (ver Tabla 45) dependiendo si alguna vez reportaron o viajaron, independiente de la semana que fuese.

Tabla 45: Clasificación de usuarios según acciones durante su período de actividad.

| Categoría | Cantidad | Porcentaje | Acumulado |
|---------------|----------------|----------------|-----------|
| Reportador | 26.844 | 17,62% | 17,62% |
| Viajador | 4.990 | 3,28% | 20,90% |
| Ambos | 13.656 | 8,97% | 29,87% |
| Sólo consulta | 106.822 | 70,13% | 100,00% |
| Total | 152.312 | 100,00% | - |

En la Tabla 46 se adjunta las estadísticas del modelo de análisis de sobrevivencia según la categoría de usuario. Se tiene que la tasa de incidencia es más alta en los usuarios que sólo consultan, y que además son los usuarios cuya semana donde el percentil 75 deja de ser activo es el más bajo, en la 14. En cambio, este valor es 37 para los reportadores, 28 para los viajeros y 41 para quienes realizan ambas acciones.

Tabla 46: Resumen estadísticas de modelación tiempo de sobrevivencia, según categoría de usuario según acciones durante su período de actividad.

| Categoría | Tiempo en riesgo | Tasa de incidencia | Tiempo de sobrevivencia | | |
|---------------|------------------|--------------------|-------------------------|-----|-----|
| | | | 25% | 50% | 75% |
| Reportador | 426.698 | 0,0444 | 3 | 14 | 37 |
| Viajador | 64.953 | 0,0591 | 1 | 8 | 28 |
| Ambos | 256.548 | 0,0386 | 4 | 17 | 41 |
| Sólo consulta | 817.716 | 0,1068 | 1 | 2 | 14 |

En la Ilustración 36, se muestra las probabilidades de sobrevivencia según categoría de usuario, en donde se ve cómo quienes sólo consultan tienen una menor probabilidad de sobrevivencia, y quienes realizan aunque sea 1 acción diferente a sólo consultar 1 vez en su período de actividad, tienen una mayor probabilidad de sobrevivencia. Dado que el análisis anterior es visual, se realiza el test log-rank para testear la igualdad de las funciones de sobrevivencia de los usuarios. Los resultados están en la tabla posterior al gráfico.

Ilustración 36: Probabilidad de supervivencia de Kaplan-Meier según categoría de usuarios en base a acciones durante su período de actividad.

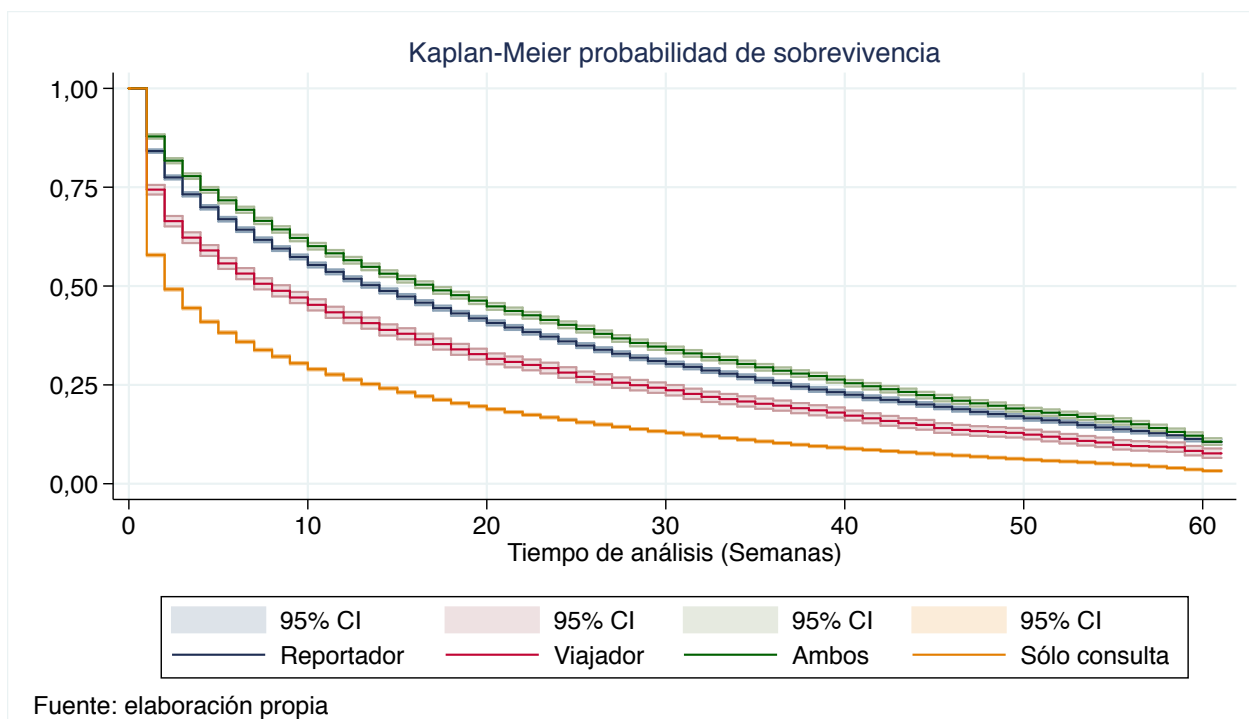


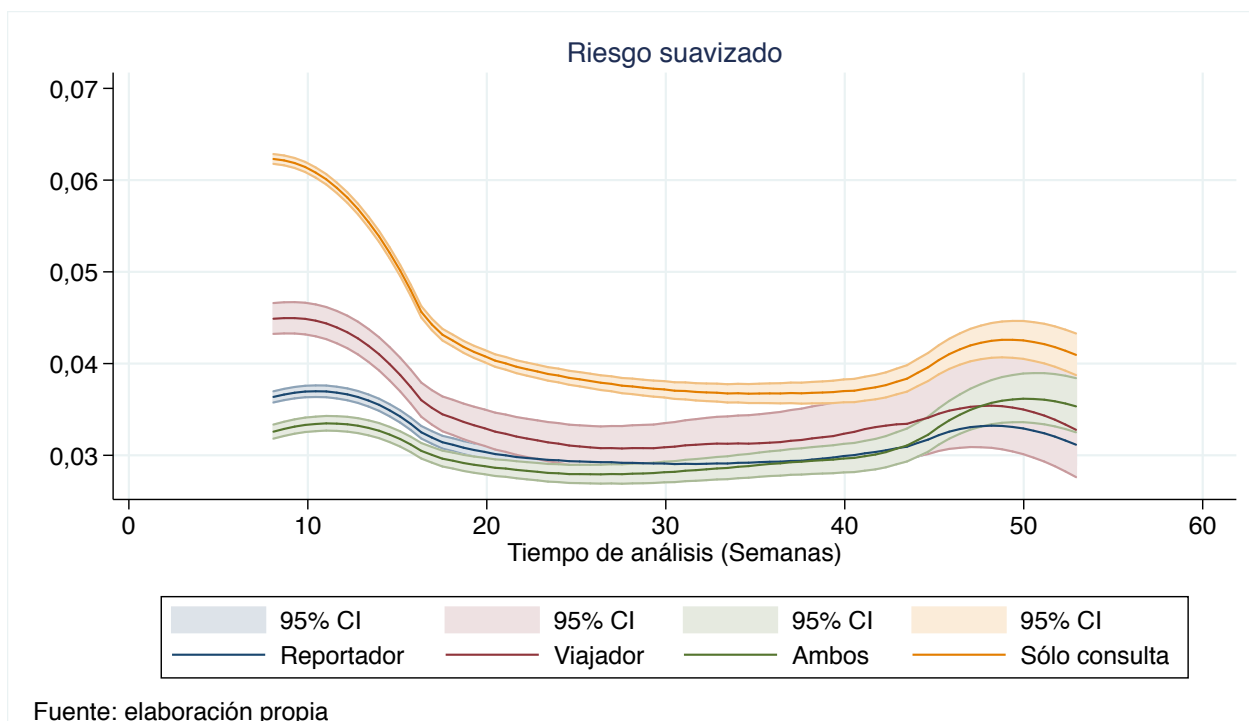
Tabla 47: Test log-rank de la igualdad de las funciones de supervivencia según categoría de usuarios según acciones durante su período de actividad.

| Categoría | Eventos observados | Eventos esperados |
|---------------|--------------------|-------------------|
| Reportador | 18.982 | 27.706,15 |
| Viajador | 3.839 | 4.536,07 |
| Ambos | 9.924 | 15.644,26 |
| Sólo consulta | 87.380 | 72.238,53 |
| Total | 120.125 | 120.125,00 |

| Test | |
|----------------|-----------|
| $\chi^2(3)$ | 10.205,43 |
| $\Pr(>\chi^2)$ | 0,000 |

Dado los resultados del test log-rank (ver Tabla 47), se rechaza la hipótesis nula, por lo que las funciones de supervivencia tienen diferencias estadísticas significativas. Para complementar lo anterior, en la Ilustración 37 se adjunta el riesgo suavizado experimentado por cada tipo de usuario, en donde se observa cómo quienes sólo consultan en la aplicación, tienen hasta la semana 40 aproximadamente, un riesgo de dejar de utilizar la aplicación más elevado que todas las otras categorías de usuarios. En cambio, dentro de los usuarios que participan de las otras funcionalidades de la aplicación, se tiene que hasta la semana 20, aproximadamente, hay diferencias entre ellos, siendo los usuarios que consultan y reportan los con menor riesgo de dejar de seguir activos consultando. Posterior a esta semana, la diferencia deja de ser significativa.

Ilustración 37: Riesgo suavizado de las categorías de usuario según las actividades realizadas durante su período de actividad.



Se procede a realizar una regresión Cox para así estimar los efectos proporcionales del ratio de riesgo, con la variable independiente los grupos de la categoría previa y grupo base quienes sólo consultan. Los resultados se adjuntan en la Tabla 48.

Tabla 48: Resultados regresión Cox.

| Categoría | Ratio de riesgo | Error estándar | Valor z | Pr(> z) |
|------------|-----------------|----------------|---------|----------|
| Reportador | 0,5541 | 0,0044 | -72,95 | 0,000 |
| Viajador | 0,6883 | 0,0113 | -22,61 | 0,000 |
| Ambos | 0,5103 | 0,005 | -62,91 | 0,000 |

Luego, el ratio de riesgo disminuye en todas las categorías de usuario respecto a los usuarios que sólo consultan en la aplicación (dada la magnitud del coeficiente, todos menor a 1, y su significancia al 1%). Se desprende así que quienes reportan y consultan, disminuyen un 44,59% el riesgo de dejar de ser activo; para quienes viajan y consultan, esta cifra es de 31,17%; y para quienes consultan, viajan y reportan, su riesgo disminuye un 48,97% de dejar de ser activo. Esto en línea con las tasas de sobrevivencia (menor tasa de riesgo, mayor probabilidad de sobrevivencia) observadas en la Ilustración 36.

4.3 Conclusiones del análisis de sobrevivencia

Es importante para la aplicación contar con usuarios activos, por lo cual conocer la tasa de retención de éstos y cómo esta varía dependiendo de otros factores es crítico para poder realizar políticas y estrategias de marketing para que éstos utilicen la aplicación activamente, y deseablemente, no sólo consulten para su utilidad propia inmediata, sino que también participen de las otras herramientas colaborativas que ayudan a *Transapp* y al resto de los viajeros.

Al considerar la actividad de la primera semana en la aplicación, quienes sólo consultan tienen una mayor probabilidad de dejar de utilizar la aplicación que quienes experimentan las otras funcionalidades. Es posible que aquellos usuarios que participan de las otras funcionalidades colaborativas sean potencialmente diferentes a quienes sólo consultan (en esta primera semana), la que puede ejemplificarse en usuarios con mayor facilidad en la tecnología (mayor navegación e interés en la aplicación), usuarios que quizás se unen a la aplicación por una cenéfa que invita a consultar paraderos versus otros que se unen por reportar problemas en el bus (expectativa objetivo de la instalación). No obstante, se recomienda realizar tutoriales más profundos a los usuarios nuevos de cómo reportar/viajar para así que comprendan a cabalidad las herramientas de la aplicación, incluso podría motivarse creándoles “misiones” de *gamificación* de bienvenida; y así comprendan la utilidad de la aplicación y tengan menor probabilidad de dejar de usar la aplicación.

Se recomienda testear estas políticas mediante experimentos aleatorios controlados para así evidenciar si esta guía efectivamente provoca una disminución en la inactividad de usuarios en el corto/mediano plazo. Por otro lado, también perfeccionar el análisis utilizando como variable de falla e inactividad, la desinstalación de la aplicación.

Finalmente, también es interesante investigar cómo motivar una mayor actividad de los usuarios en las diferentes herramientas de la aplicación, lo cual podría potencialmente disminuir el riesgo de convertirse en inactivo, por lo cual en el siguiente capítulo se realiza un experimento de campo aleatorio que estudia cómo potenciar los usuarios reportadores en *Transapp*.

Capítulo 5

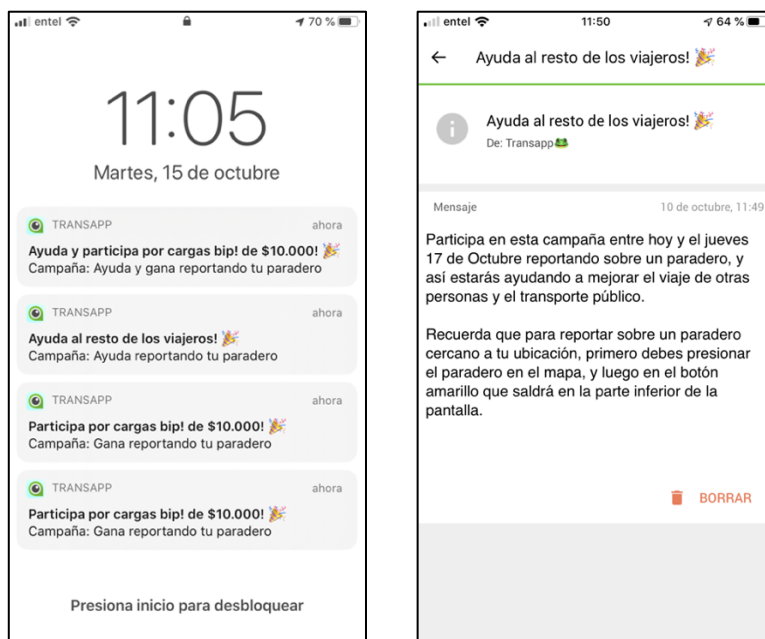
Experimento de campo

En este capítulo se describe un experimento de campo realizado (como parte de este trabajo de título) en la aplicación *Transapp*, con el objetivo de investigar cuál incentivo, si monetario o prosocial o una combinación de ambos, es el más efectivo para lograr que los usuarios reporten en la aplicación, y así aumentar la participación en la aplicación.

5.1 Metodología

El experimento se realiza en *Transapp* mediante envío de notificaciones *push*²⁰ a los celulares, las cuales además se acumulan en la bandeja de entrada de la aplicación junto a las otras notificaciones recibidas anteriormente. Las notificaciones se componen de tres apartados en su contenido: título, subtítulo y contenido. En la Ilustración 38 se observa las notificaciones *push*, donde el título se encuentra destacado en negrita y el subtítulo está inferior a éste. Por otro lado, el contenido sólo se muestra cuando se abre la notificación o se accede a la bandeja de entrada en la misma aplicación.

Ilustración 38: Ejemplo de recepción de todas las notificación *push* enviadas y la apertura de un mensaje en la bandeja de entrada de los usuarios.



Fuente: obtenidas de la aplicación *Transapp* (2019).

²⁰ Las notificaciones *push* son mensajes que se de una aplicación fuera del ambiente de ésta (Google, 2019).

5.2 Condiciones experimentales

Debido al funcionamiento de la aplicación para el envío de mensajes, éstos deben tener una ID destinataria definida, por lo que se selecciona todos los usuarios que utilizaron *efectivamente* la aplicación durante Septiembre de 2019, es decir, que consultaron al menos una vez el tiempo de llegada de los buses en algún paradero. Estos totalizaron 46.516²¹ usuarios activos (consultores), a quienes se asigna aleatoriamente a una de las siguientes condiciones experimentales:

1. **Incentivo económico:** motiva a las personas a reportar mediante la participación en un sorteo por tres cargas bip!²² de \$10.000. A continuación, se referirá a ésta como condición experimental “Económico”.
2. **Incentivo económico – con texto prosocial:** es una variante de la condición previa, ya que se diseña utilizando el mismo título y subtítulo de notificación, sólo diferenciando el contenido del mensaje que fue un incentivo prosocial y económico (el mismo utilizado en la condición experimental 4). A continuación, se referirá a ésta como condición experimental “Económico - texto Prosocial”²³.
3. **Incentivo prosocial:** motiva a las personas destacando la utilidad del reporte hacia al resto de las personas y el transporte público. A continuación, se referirá a ésta como condición experimental “Prosocial”.
4. **Incentivo económico y prosocial:** es una combinación del incentivo económico y prosocial, tanto en el título de la notificación como en el subtítulo. A continuación, se referirá a ésta como “Ambos” (aludiendo a ambos incentivos).
5. **Control:** no reciben notificación.

El experimento se realiza durante 3 días²⁴ consecutivos, en los cuales se envía tres notificaciones *push* a los usuarios, una diaria. Estas notificaciones se enmarcan en una campaña denominada “Ayuda y/o gana reportando sobre tu paradero” (dependiendo de la condición experimental).

La hora de recepción de la notificación se programa en función de la hora en la cual los usuarios consultaron una mayor cantidad de veces el tiempo de espera buses en el mes de Septiembre, así intentando acercarse a la hora en la que era más probable que las personas estuviesen más próximas a utilizar la aplicación asociado a un hábito de viaje.

²¹ Si bien los 46.516 no son exactos a los 47.320 desplegados en la Tabla 6, la diferencia se debe a los usuarios que fueron partícipes de un testeo del experimento, realizado previamente con cantidad limitada de usuarios.

²² La tarjeta bip! es el medio de pago del sistema de transporte público de Santiago (Tarjeta bip!, 2019). Funciona con carga monetaria de prepago por lo que el premio equivale a 13,3 viajes en horario valle (valor en función de pasaje en 2019).

²³ En ciertas tablas y figuras, por temas de espacio, se abreviará como “Econ. – txt. Pros.”.

²⁴ Se utilizó este plazo debido a que en una encuesta previa realizada por *Transapp*, el 98% de las respuestas fueron obtenidas en aquel intervalo de tiempo, desde la primera notificación.

En la Tabla 49 se muestra las configuraciones de los apartados de las diferentes condiciones experimentales (los mensajes de los otros días se adjuntan en Anexos, Apéndice K: Condiciones experimentales). En particular, se observa la igualdad en título y subtítulo de las condiciones “Económico” y “Económico – texto Prosocial”, y la diferencia de éstos en el mensaje principal, en donde el de “Económico – texto Prosocial” es el mismo que en “Ambos”. Por otro lado, notar que todos los mensajes incluyen un párrafo en común de cómo reportar en la aplicación, esto debido a que la acción de reportar es lejana a la acción de revisar una notificación en la bandeja de entrada y así también poder guiar a quienes no supiesen cómo hacerlo.

Tabla 49: Configuración y detalle de la notificación de cada condición experimental del día 1.

| Apartado | “Económico” | “Económico – texto Prosocial” | “Prosocial” | “Ambos” |
|---|--|--|--|--|
| Notificación (título y subtítulo) | Participa por cargas bip! de \$10.000! 🎉 | Participa por cargas bip! de \$10.000! 🎉 | Ayuda al resto de los viajeros! 🎉 | Ayuda y participa por cargas bip! de \$10.000! 🎉 |
| | Campaña: Gana reportando tu paradero | Campaña: Gana reportando tu paradero | Campaña: Ayuda reportando tu paradero | Campaña: Ayuda y gana reportando tu paradero |
| Mensaje particular de cada condición experimental | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y estarás participando en un sorteo por tres cargas bip! de \$10.000. | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de \$10.000. | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de \$10.000. |
| Mensaje común para todas las condiciones experimentales | <p>Recuerda que para reportar sobre un paradero cercano a tu ubicación, primero debes presionar el paradero en el mapa, y luego en el botón amarillo que saldrá en la parte inferior de la pantalla.</p> <p>Gracias por ser parte de nuestra comunidad 🍌</p> | | | |

5.3 Descripción de los participantes

Los 46.516 usuarios activos seleccionados, además se clasifican en base a su participación en la aplicación en la actividad de reportes realizados, tanto de buses como de paraderos, durante los meses de Agosto y Septiembre de 2019²⁵. El criterio de clasificación (ver Tabla 50) toma en consideración separar aquellos usuarios *activos* de los *inactivos* (en la actividad de reportes), y luego asemejar una proporción a la regla de Nielsen así obteniendo un nuevo grupo de usuarios activos pero en menor medida, que se han denominado “Reportadores Tímidos”.

Tabla 50: Criterio de clasificación de usuarios activos para realización de experimento.

| Clasificación | Número de reportes realizados |
|------------------------|-------------------------------|
| “Reportadores” | Más de 10 |
| “Reportadores Tímidos” | Entre 1 y 9 |
| “No Reportadores” | Cero |

Las estadísticas de usuarios en base a la clasificación previa, se observa en la Tabla 51. La proporción final de usuarios utilizada fue de 82,5:15:2,5, similar a la 90:9:1 elaborada por Nielsen.

Tabla 51: Estadísticas de clasificación de usuarios según reportes realizados.

| Clasificación | Frecuencia | Porcentaje | Acumulado |
|------------------------|---------------|---------------|-----------|
| “No Reportadores” | 38.380 | 82,5% | 82,5% |
| “Reportadores Tímidos” | 6.995 | 15,0% | 97,5% |
| “Reportadores” | 1.141 | 2,5% | 100,0% |
| Total | 46.516 | 100,0% | - |

Finalmente los usuarios se asignan a las diferentes condiciones experimentales utilizando una aleatorización en bloque, en base a la clasificación previa según tipo de usuario, cuya distribución se muestra en la Tabla 52. La proporción utilizada es la que se observa en la columna “Porcentaje” de la Tabla 53.

²⁵ Se destaca que los reportes utilizados para esta determinación no fueron filtrados como se realiza en el capítulo 3.

Tabla 52: Distribución de usuarios en las condiciones experimentales según clasificación de uso en la aplicación

| Condición experimental | Reportadores | Rep. Tímidos | No Rep. | Total |
|------------------------|--------------|--------------|---------------|---------------|
| “Económico” | 274 | 1.679 | 9.211 | 11.164 |
| “Econ. – txt. Pros.” | 274 | 1.679 | 9.211 | 11.164 |
| “Prosocial” | 160 | 979 | 5.373 | 6.512 |
| “Ambos” | 274 | 1.679 | 9.211 | 11.164 |
| “Control” | 159 | 979 | 5.374 | 6.512 |
| Total | 1.141 | 6.995 | 38.380 | 46.516 |

Las proporciones de las diferentes condiciones experimentales se define con el objetivo de incrementar el poder estadístico, así aumentando la muestra en aquellas condiciones experimentales en las cuales la magnitud del efecto esperado de la notificación²⁶ fuese similar entre ellas, ejemplo de esto son: “Económico”, “Económico – texto Prosocial” y “Ambos”, en donde para identificar un efecto diferenciador entre ellas, se requiere mayor muestra así *sacrificando* menor cantidad de usuarios en las otras condiciones experimentales, “Prosocial” y “Control”, las cuales tienen un efecto esperado menor (proporción de personas que reportaron).

Tabla 53: Distribución de usuarios en las diferentes condiciones experimentales.

| Condición experimental | Frecuencia | Porcentaje | Acumulado |
|------------------------|---------------|---------------|-----------|
| “Económico” | 11.164 | 24,0% | 24,0% |
| “Econ. – txt. Pros.” | 11.164 | 24,0% | 48,0% |
| “Prosocial” | 6.512 | 14,0% | 62,0% |
| “Ambos” | 11.164 | 24,0% | 86,0% |
| “Control” | 6.512 | 14,0% | 100,0% |
| Total | 46.516 | 100,0% | - |

²⁶ Resultados obtenidos en base a realización de un testeo del experimento en la semana previa del experimento final.

5.4 Exposición al tratamiento

En este experimento el tratamiento “completo” o intensivo, se considera como la lectura de la notificación y el contenido del mensaje, por lo cual quienes reciben la notificación en sus celulares y no abren la notificación, reciben un tratamiento “limitado” o extensivo dado que sólo se ven expuestos parcialmente al tratamiento que sería la lectura de la notificación (se considera supuesto la lectura).

La Tabla 54 considera aspectos generales de la asignación, recepción y apertura de las notificaciones, y la cantidad de reportes realizados por condición experimental.

Tabla 54: Estadísticas sobre recepción y exposición a los tratamientos de las condiciones experimentales.

| Condición experimental | Usuarios asignados | Reciben al menos una notificación | Abren al menos una notificación | Reportan al menos una vez ²⁷ |
|------------------------|--------------------|-------------------------------------|---------------------------------|---|
| “Control” | 6.512 | - | - | 44 (0,68%) |
| “Económico” | 11.164 | 4.223 (37,83%) | 1.049 (24,84%) | 405 (10,96%) |
| “Econ. – txt. Pros.” | 11.164 | 4.269 (38,24%) | 1.014 (23,75%) | 364 (9,60%) |
| “Prosocial” | 6.512 | 2.439 (37,45%) | 286 (11,73%) | 64 (3,36%) |
| “Ambos” | 11.164 | 4.234 (37,93%) | 829 (19,58%) | 265 (7,20%) |
| Total | 46.516 | 15.165 (37,91%)²⁸ | 3.178 (20,96%) | 1.304 (6,02%)²⁹ |

Quienes reciben el tratamiento extensivo generan un problema denominado *One-Sided Non-Compliance* (en español, “No cumplimiento por un lado”) en el cual las personas asignadas al tratamiento intensivo deciden si tomarlo o no, que en este caso sería abrir la notificación, ya sea por un error o una acción deliberada.

A partir de la Tabla 54 se tiene que la tasa de apertura de las notificaciones es de 20,96%, aunque varía dependiendo de la condición experimental, la más baja siendo 11,73% en la condición “Prosocial” y la más alta con 24,84% en la condición “Económico” (las diferencias entre ellas se aborda a mayor profundidad en el análisis según apertura).

Por otro lado, la exposición al tratamiento claramente primero requiere la recepción de la notificación, no obstante, debido a un problema en el envío de las notificaciones en la aplicación, sólo el 37,91% de los usuarios recibe al menos una de las notificaciones (más detalle en Tabla 54).

²⁷ Porcentaje respecto al número de personas que reciben al menos una notificación.

²⁸ Respecto al número total de usuarios asignados a un grupo experimental diferente al de control, que totalizan 40.004 usuarios.

²⁹ Respecto al número de usuarios que recibieron al menos una notificación y los usuarios del grupo de control, que totalizan 21.677 usuarios.

Las razones específicas de este problema son desconocidos para el equipo de la aplicación³⁰, el cual fue identificado en un testeo previo del experimento, por lo que para disminuir la tasa de no envío se aumenta el tiempo en las cuales era posible que los usuarios recibieran la notificación (de ser una notificación inmediata a tener un tiempo de vida de 3 horas desde la hora programada³¹). A pesar de esto, la tasa de recepción no aumentó de gran manera, ligeramente de un aproximado 34% a un 37%.

Para estudiar si la tasa de recepción de las diferentes notificaciones varía según condición experimental, se realiza una estimación de un modelo logit de la variable dependiente “Recepción” (de al menos una notificación) controlando por las diferentes notificaciones como variables independientes, con “Ambos” como intercepto. Los resultados se muestran en la Tabla 55.

Tabla 55: Resultado regresión logit, estudio de diferencias de tasa de recepción entre notificaciones.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| “Económicos” | -0,004 | 0,027 | -0,15 | 0,879 |
| “Ec. – txt. Pr.” | 0,013 | 0,027 | 0,48 | 0,630 |
| “Prosocial” | -0,020 | 0,032 | -0,62 | 0,533 |
| Intercepto | -4,927 | 0,019 | -25,26 | 0,000 |

Al ser ningún coeficiente de la regresión significativo, se desprende que la tasa de recepción de las notificaciones no son significativamente diferentes respecto a la de “Ambos” (intercepto). Al realizar una comparación por pares (ver Tabla 56) se concluye que ninguna tiene diferencias estadísticas entre sí. Por lo que efectivamente no hay diferencia estadística en la tasa de recepción de las notificaciones entre las condiciones experimentales.

Tabla 56: Comparación por pares entre notificaciones para identificar diferencias en tasa de recepción.

| Condiciones experimentales | Contraste | Valor z | Pr(> z) |
|-----------------------------------|------------------|----------------|--------------------|
| “Económico” vs. “Ec. – txt. Pr.” | -0,017 | -0,63 | 0,526 |
| “Económico” vs. “Prosocial” | 0,015 | 0,49 | 0,622 |
| “Económico” vs “Ambos” | -0,004 | -0,15 | 0,879 |
| “Ec. – txt. Pr.” vs. “Prosocial” | 0,033 | 1,04 | 0,300 |
| “Ec. – txt. Pr.” vs. “Ambos” | 0,013 | 0,48 | 0,630 |
| “Prosocial” vs. “Ambos” | -0,020 | -0,62 | 0,533 |

³⁰ El problema radica en el funcionamiento de *Reports Api* (de Google) que es la herramienta que envía las notificaciones a los usuarios.

³¹ Por ejemplo, si un usuario fue asignado a recibir la notificación a las 10:00 am., si no tenía internet a esa hora no recibiría la notificación, por lo que al aumentar el intervalo de recepción hasta las 13:00 pm. (+3 horas), podría recibir la notificación en el primer momento que tuviese internet dentro de aquel intervalo.

El análisis de resultados del experimento omite aquellos usuarios que estaban asignados a recibir una notificación y no lo hicieron, los cuales totalizan el 62,09%. Esto no es un problema ya que al no recibir una notificación debiesen “comportarse” como el grupo de control (en términos de su probabilidad de emitir un reporte) lo que se comprueba en Anexos, Apéndice L: Análisis de resultados de experimento considerando usuarios que no reciben las notificaciones, donde se demuestra que no hay diferencias estadísticas significativas entre las diferentes condiciones experimentales y el grupo de control.

Es importante mencionar que durante el segundo día del experimento, el equipo de *Transapp* emitió dos notificaciones no planificadas sobre interrupciones del servicio de Metro de Santiago a todos los usuarios, lo que dificultó la obtención de datos claros sobre cuáles usuarios recibieron cuál notificación, si ya hubiese la del experimento y/o la de las interrupciones, ya que la base de datos no discrimina por nombre de notificación, sino que por evento de recepción y apertura según hora. Para disminuir el impacto de este problema, se realizó un trabajo de filtro para asegurar la disminución de este error a una mínima expresión, en el cual se trabajó con diferentes condiciones de recepción de mensajes según horario de llegada de la notificación y la hora asignada de la llegada de cada notificación a cada usuario. Lo mismo se realizó para la apertura de notificaciones. El análisis realizado en el siguiente capítulo se considera correcto, aunque de todas formas en Anexos, Apéndice M: Análisis de resultados omitiendo segundo día, se realiza un nuevo análisis de resultados omitiendo todas las recepciones y aperturas de la notificación durante el segundo día, los cuales son bastante similares a los obtenidos en el análisis final (utilizando la información de todos los días a partir del trabajo de filtro realizado), que se describe en el siguiente apartado.

5.5 Análisis de resultados

El experimento consta de dos tipos de análisis de sus tratamientos, uno extensivo y uno intensivo. Debido a la diferencias en la tasa de apertura de las notificación, se concluye que el título y subtítulo de la notificación provocaron diferentes intereses en los individuos que la recibieron, siendo aquellos que involucraban un incentivo económico provocaron una mayor atención, generando así un sesgo de selección de quienes finalmente abren la notificación y exponiéndose así al tratamiento intensivo. Esto significa que una comparación de los grupos que abrieron las notificaciones entre sí no es válida, porque es muy poco probable que sean individuos con características similares.

Por lo anterior, se procede a realizar el análisis de resultados en dos partes: primero se analiza bajo la recepción de la notificación, y luego se analiza considerando la apertura de la notificación. Para la recepción de la notificación, se analiza el efecto “*Intention to Treat*” (intención del tratamiento), abreviado “ITT”, que simboliza el efecto causal de la recepción de la notificación. Para el caso de la apertura, se determina el efecto “*Complier Average Causal Effect*” (efecto promedio causal de los cumplidores), abreviado “CACE”, que simboliza el efecto causal de la apertura de la notificación en los individuos que efectivamente la abren.

En ambos enfoques de análisis, tanto de recepción como apertura, se realiza dos modelos para estimar los efectos. El primero es el “modelo simple”, que contendrá sólo las condiciones experimentales, y el segundo es el “modelo por tipo de usuario”, que agrega heterogeneidad a las regresiones dependiendo de la clasificación de reportador (ver Tabla 50).

5.5.1 Análisis de resultados – según recepción

Al analizar el experimento dada la recepción de la notificación, sin importar su apertura, se procede a unir las notificaciones “Económico” y “Económico – texto Prosocial” ya que éstas son idénticas en título y subtítulo (que es lo que se observaba al momento de la recepción). Su unión se denominará como “Económicos”.

5.5.1.1 Análisis – modelo simple

En la Tabla 57 se observa la cantidad de usuarios asignados a cada una de las condiciones experimentales, así como también cuántos recibieron la notificación y cuántos de éstos reportan sobre un paradero.

Tabla 57: Estadísticas para estimación del ITT - modelo simple.

| Condición experimental | Usuarios asignados | Reciben al menos una notificación | Reportan al menos una vez | ITT |
|------------------------|--------------------|-----------------------------------|---------------------------|-------|
| “Económicos” | 22.328 | 8.492 (38,03%) | 769 (9,06%) | 8,38% |
| “Prosocial” | 6.512 | 2.439 (37,45%) | 64 (2,62%) | 1,94% |
| “Ambos” | 11.164 | 4.234 (37,93%) | 265 (6,26%) | 5,58% |
| “Control” | 6.512 | - | 44 (0,68%) | - |

En la última columna se encuentra el “*Intention to Treat*” que es el efecto de la recepción de la notificación en la probabilidad de emitir un reporte (este omite la apertura de la notificación). Este se calcula como la diferencia en la proporción de usuarios que reportaron en cada una de las condiciones experimentales respecto al grupo de control (que no recibieron notificación). Por ejemplo, la proporción de usuarios que reportaron dado que recibieron la notificación “Económico” es de 9,06%, en contraste al 0,68% del grupo de control, cuya diferencia entre ambos es de 8,38%. Esto implica que la recepción de una notificación con incentivo económico aumentaría un 8,38% la probabilidad de emitir un reporte respecto a la no recepción de alguna (y así para el resto de las condiciones experimentales).

Este primer análisis omite la significancia estadística de las notificaciones respecto a la no recepción y entre ellas, por lo que a continuación se procede a estimar el siguiente modelo mediante una regresión logística (ver Ecuación 3).

Ecuación 3: Especificación del modelo logit - simple, para evaluación del experimento, enfoque ITT.

$$Pr(X_i) = B_0 + \sum_{j=\{Econs,Prs,Amb\}} B_j * Condición_{i,j} + e_i$$

En donde, se considera:

- X_i : variable binaria que toma valores 1 o 0, dependiendo si la persona i reportó sobre un paradero.
- $Condición_{i,j}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i fue asignada a la condición experimental j , las cuales son “Económicos”, “Prosocial” o “Ambos”.

Los resultados de este modelo logit se muestran en la Tabla 58, junto a los *Odds-ratio* y los efectos marginales promedios respecto al grupo de control. Además se tiene que la estimación entrega prácticamente los mismos valores que al calcular el ITT *manualmente*.

Tabla 58: Resultados modelo logit - simple, para estimación del efecto ITT.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) | Odds-ratio | Efecto marginal³² |
|---------------------|------------------|-----------------------|----------------|--------------------|-------------------|-------------------------------------|
| “Económicos” | 2,684 | 0,156 | 17,21 | 0,000 | 14,637 | 8,38% |
| “Prosocial” | 1,377 | 0,197 | 6,98 | 0,000 | 3,961 | 1,95% |
| “Ambos” | 2,284 | 0,164 | 13,92 | 0,000 | 9,814 | 5,58% |
| Intercepto | -4,990 | 0,151 | -32,99 | 0,000 | 0,006 | - |

Se observa que todos los coeficientes son significativos y positivos, por lo que todas las condiciones experimentales aumentan la probabilidad de emitir un reporte (respecto a la no recepción de alguna). Al analizar según efectos marginales promedios, la notificación “Económicos”, que es la que tiene mayor efecto, su recepción aumenta en un 8,38% la probabilidad de emitir un reporte sobre un paradero respecto al grupo de control.

Respecto a la diferencia entre notificaciones, se procede a realizar una “*pairwise comparison*”³³ o comparación por pares, para ver diferencias estadísticas significativas entre ellas, cuyos resultados se observan en la Tabla 59.

Tabla 59: Resultados análisis comparación por pares entre notificaciones, del modelo simple en efecto ITT.

| Condiciones experimentales | Contraste | Valor z | Pr(> z) |
|-----------------------------------|------------------|----------------|--------------------|
| “Económicos” vs. “Prosocial” | 1,307 | 9,89 | 0,000 |
| “Económicos” vs. “Ambos” | 0,399 | 5,41 | 0,000 |
| “Económicos” vs. “Control” | 2,683 | 17,21 | 0,000 |
| “Prosocial” vs “Ambos” | -0,907 | -6,40 | 0,000 |
| “Prosocial” vs. “Control” | 1,376 | 6,98 | 0,000 |
| “Ambos” vs. Control” | 2,283 | 13,92 | 0,000 |

Se concluye que todas las notificaciones tienen diferencias estadísticas significativas entre sí, por lo que todas las notificaciones aumentan la probabilidad de emitir un reporte y este efecto es diferente entre cada una de las notificaciones, siendo las que involucran un incentivo meramente económico más efectivas que aquellas que tienen una componente prosocial.

³² Efecto marginal promedio (*Average marginal effect*, “AME”) respecto al grupo control, medido en probabilidad de emitir reporte. Resultado obtenido utilizando comando “*margins, dydx*” en *Stata*.

³³ Realizada utilizando el comando “*pwcompare*” en *Stata*.

5.5.1.2 Análisis - modelo según tipo de usuario

Tal como se indicó en la sección “Descripción de los participantes”, el experimento consideró asignación aleatoria en bloque de los tres tipos de usuarios identificados para así poder determinar heterogeneidad de los diferentes tratamientos.

El cálculo *manual* del efecto “ITT” es prácticamente idéntico al obtenido del siguiente modelo econométrico, por lo que se omite su cálculo en tablas (como en la Tabla 57Tabla 57).

Para el análisis se utilizará un modelo logit, cuya especificación es la siguiente (ver Ecuación 4).

Ecuación 4: Especificación del modelo logit – tipo de usuario, para evaluación del experimento, enfoque ITT.

$$Pr(X_i) = B_0 + \sum_{j=\{Ecs,Prs,Amb\}} \sum_{k=\{NoRep,RepTim,Rep\}} B_j * Condición_{i,j} * TipoUsuario_{i,k} + e_i$$

Las variables de la ecuación se definen a continuación como:

- X_i : variable binaria que toma valores 1 o 0, dependiendo si la persona i reportó sobre un paradero.
- $Condición_{i,j}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i fue asignada a la condición experimental j , las cuales son “Económicos”, “Prosocial” o “Ambos”.
- $TipoUsuario_{i,k}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i es usuario tipo “No Reportador”, “Reportador Tímido” o “Reportador”.

Los resultados de este modelo logit se muestran en la Tabla 60, junto a los *Odds-ratio* y los efectos marginales promedios respecto al grupo de control de los usuarios “No Reportadores”.

Tabla 60: Resultados modelo logit - tipo de usuario, para estimación del efecto ITT.

| Coeficiente | | Estimador | Error estándar | Valor Z | Pr(> z) | Odds-ratio | Efecto marginal ³⁴ |
|---------------------|--------------|-----------|----------------|---------|----------|------------|-------------------------------|
| “Reportador” | “Económicos” | 5,376 | 0,299 | 17,96 | 0,000 | 216,28 | 35,83% |
| | “Prosocial” | 5,453 | 0,368 | 14,79 | 0,000 | 233,45 | 37,61% |
| | “Ambos” | 5,052 | 0,346 | 14,57 | 0,000 | 156,38 | 28,73% |
| | “Control” | 3,757 | 0,375 | 10,00 | 0,000 | 42,83 | 9,80% |
| “Reportador Tímido” | “Económicos” | 4,239 | 0,278 | 15,25 | 0,000 | 69,36 | 15,07% |
| | “Prosocial” | 2,531 | 0,389 | 6,51 | 0,000 | 12,56 | 2,91% |
| | “Ambos” | 4,117 | 0,291 | 14,14 | 0,000 | 61,36 | 13,55% |
| | “Control” | 1,714 | 0,380 | 4,52 | 0,000 | 5,55 | 1,16% |
| “No Reportador” | “Económicos” | 3,341 | 0,271 | 12,29 | 0,000 | 28,26 | 6,61% |
| | “Prosocial” | 1,636 | 0,332 | 4,92 | 0,000 | 5,13 | 1,06% |
| | “Ambos” | 2,828 | 0,280 | 10,08 | 0,000 | 16,91 | 3,97% |
| Intercepto | | -5,947 | 0,276 | -22,23 | 0,000 | 0,002 | - |

En la tabla anterior se observa que todas los coeficientes asociados a notificaciones son positivos y significativos, por lo que la recepción de una notificación aumenta la probabilidad de emitir un reporte respecto al grupo de control de la categoría “No Reportador”. Esta magnitud depende del tipo de usuario, por lo cual se procede a realizar un cambio de base en la Tabla 61 para así aislar el efecto ITT.

Tabla 61: Resultados de efecto ITT a partir de modelo - tipo de usuario.

| Condición experimental | Reportadores | Reportadores tímidos | No Reportadores |
|------------------------|--------------|----------------------|-----------------|
| “Económicos” | 26,03% | 13,91% | 6,61% |
| “Prosocial” | 27,81% | 1,75% | 1,06% |
| “Ambos” | 18,93% | 12,39% | 3,97% |
| “Control” | Base | Base | Base |

Se tiene así que el efecto de la recepción de la notificación varía dependiendo de la clasificación de usuario de gran manera. Por ejemplo, las notificaciones tienen un efecto mayor a medida que los usuarios son más activos en la aplicación, no obstante, una importante detalle es que en los usuarios “Reportadores Tímidos” y “No Reportadores” la notificación “Prosocial” es aquella con menor efecto, no así en los usuarios “Reportadores” en donde es la mayor por un margen pequeño.

³⁴ Efecto marginal promedio (*Average marginal effect*, AME) respecto al grupo “Control – No Reportador”, medido en probabilidad de emitir reporte. Resultado obtenido utilizando comando “*margins, dydx*” en *Stata*.

Analizada la significancia respecto al grupo de control, se procede a testear la diferencia entre notificaciones usando comparaciones por pares, cuyos resultados se encuentran en la Tabla 62.

Tabla 62: Resultados análisis comparación por pares entre notificaciones, del modelo tipo de usuario en efecto ITT

| Tipo de usuario | Condiciones experimentales | Contraste | Valor t | Pr(> t) |
|---------------------|------------------------------|-----------|---------|----------|
| "Reportador" | "Económicos" vs. "Prosocial" | -0,076 | -0,27 | 0,790 |
| | "Económicos" vs. "Ambos" | 0,324 | 1,26 | 0,209 |
| | "Económicos" vs. "Control" | 1,619 | 5,47 | 0,000 |
| | "Prosocial" vs. "Ambos" | 0,400 | 1,19 | 0,233 |
| | "Prosocial" vs. "Control" | 1,695 | 4,63 | 0,000 |
| | "Ambos" vs. "Control" | 1,294 | 3,77 | 0,000 |
| "Reportador Tímido" | "Económicos" vs. "Prosocial" | 1,708 | 5,85 | 0,000 |
| | "Económicos" vs. "Ambos" | 0,122 | 0,89 | 0,372 |
| | "Económicos" vs. "Control" | 2,524 | 9,03 | 0,000 |
| | "Prosocial" vs. "Ambos" | -1,585 | -5,21 | 0,000 |
| | "Prosocial" vs. "Control" | 0,816 | 2,10 | 0,036 |
| | "Ambos" vs. "Control" | 2,402 | 8,21 | 0,000 |
| "No Reportador" | "Económicos" vs. "Prosocial" | 1,705 | 8,40 | 0,000 |
| | "Económicos" vs. "Ambos" | 0,513 | 5,31 | 0,000 |
| | "Económicos" vs. "Control" | 3,341 | 12,29 | 0,000 |
| | "Prosocial" vs. "Ambos" | -1,192 | -5,56 | 0,000 |
| | "Prosocial" vs. "Control" | 1,636 | 4,92 | 0,004 |
| | "Ambos" vs. "Control" | 2,828 | 10,08 | 0,000 |

Se desprende que:

- En la clasificación "Reportadores", no hay diferencias significativas entre las notificaciones, pero sí las hay respecto al grupo de control.
- Para los usuarios "Reportadores Tímidos", todas las notificaciones tienen diferencias estadísticas significativas entre sí, a excepción entre "Económicos" y "Ambos", que involucran un incentivo económico.
- Finalmente, para los usuarios "No Reportadores" todas las notificaciones tienen diferencias estadísticas significativas entre sí.

5.5.2 Análisis de resultados – según apertura

En esta sección se estima el efecto de la notificación y la apertura de ésta, utilizando el enfoque “CACE” así corrigiendo la endogeneidad de la variable apertura, producida por el sesgo de selección.

Dado que se evalúa la diferencia en la probabilidad de reportar dada la apertura de la notificación, se separa la variable “Económicos” en “Económico” y “Económico – texto Prosocial” (como en el inicio), ya que la diferencia entre ambas se observa en el contenido del mensaje, el cual se despliega al momento de abrir la notificación (que es el enfoque en el cual se analiza a continuación).

Al realizar un análisis para identificar diferencias estadísticas en la tasa de apertura de las notificaciones, se procede a estimar un modelo logit con variable dependiente “Apertura” y dependientes las diferentes condiciones experimentales, dejando como caso base “Económico – texto Prosocial”. Los resultados se muestran en la Tabla 63.

Tabla 63: Análisis modelo logit para identificar diferencias estadísticas entre tasas de apertura.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) | Efecto marginal |
|--------------|-----------|----------------|---------|----------|-----------------|
| “Económico” | 0,059 | 0,050 | 1,17 | 0,243 | 0,01% |
| “Prosocial” | -0,852 | 0,072 | -11,76 | 0,000 | -12,02% |
| “Ambos” | -0,246 | 0,053 | -4,66 | 0,000 | -4,17% |
| Intercepto | -1,166 | 0,035 | -32,43 | 0,000 | - |

La notificación “Prosocial” y “Ambos” efectivamente tienen diferencias estadísticas significativas respecto a la condición base (“Económico – texto Prosocial”). En cambio, para la notificación “Económico” no existe evidencia para concluir que hay diferencia en su tasa de apertura, lo cual es esperable al tener ambas la misma notificación, dado lo anterior, por lo que se concluye que potencialmente son usuarios comparables por lo que si se identifica una diferencia en su comportamiento, se debe al mensaje de la notificación.

5.5.2.1 Análisis – modelo simple

Para calcular el “CACE”, se recurre a la ecuación:

Ecuación 5: Identidad para cálculo de efecto CACE.

$$CACE = \frac{ITT}{ITT_d}$$

En donde, ITT corresponde al efecto “Intention to Treat” e ITT_d a la proporción de usuarios cumplidores, que son quienes abren la notificación (quienes se exponen al tratamiento intensivo).

La división entre ambas cifras indica el “CACE”, cuyo cálculo se realiza en las siguientes dos tablas (ver Tabla 64 y Tabla 65), en donde se observan: la cantidad de usuarios asignados a cada una de las condiciones experimentales, así como también cuántos recibieron la notificación, cuántos de éstos la abrieron y finalmente reportaron.

Tabla 64: Estadísticas para estimación del ITT - modelo simple.

| Condición experimental | Usuarios asignados | Reciben al menos una notificación | Reportan al menos una vez | ITT |
|------------------------|--------------------|-----------------------------------|---------------------------|-------|
| “Económico” | 11.164 | 4.223 (37,83%) | 405 (9,59%) | 8,91% |
| “Econ. – txt. Pr.” | 11.164 | 4.269 (38,24%) | 364 (8,53%) | 7,85% |
| “Prosocial” | 6.512 | 2.439 (37,45%) | 64 (2,62%) | 1,94% |
| “Ambos” | 11.164 | 4.234 (37,93%) | 265 (6,26%) | 5,58% |
| “Control” | 6.512 | - | 44 (0,68%) | - |

Si bien es similar a la Tabla 57, aquella une los incentivos económicos en una condición experimental, lo cual acá no es válido, por lo que se realiza el cálculo para obtener los “ITT” de las diferentes notificaciones, cuyo resultado está en la última columna.

Tabla 65: Estadísticas para estimación del CACE - modelo simple.

| Condición experimental | Reciben al menos una notif. | Abren al menos una notif. [ITT_D] | Reportan al menos una vez | CACE |
|------------------------|-----------------------------|---------------------------------------|---------------------------|--------|
| “Económico” | 4.223 | 1.049 (24,84%) | 348 (33,17%) | 35,87% |
| “Econ. – txt. Pr.” | 4.269 | 1.014 (23,75%) | 310 (30,57%) | 33,05% |
| “Prosocial” | 2.439 | 286 (11,73%) | 38 (13,29%) | 16,54% |
| “Ambos” | 4.234 | 829 (19,58%) | 210 (25,33%) | 28,50% |
| “Control” | - | - | 44 (0,68%) | - |

En la Tabla 65, se encuentra la cantidad de usuarios que abren la notificación, es decir que se exponen al tratamiento, y la proporción de éstos, en paréntesis, simboliza el ITT_D que es crítico para el cálculo del “CACE”, que se encuentra en la última columna. Por ejemplo, para el obtener el efecto de la apertura de la notificación del incentivo “Económico” hay que dividir su “ITT”, de 8,91%, con la proporción de usuarios que abrieron aquella notificación recibida, que es el 24,84%; con lo cual se obtiene el 35,87%. Esta cifra significa que quienes abren la notificación “Económico” aumentan un 35,87% su probabilidad de emitir un reporte.

El anterior análisis omite la significancia estadística de las notificaciones, por lo que se procede a estimar un modelo utilizando una regresión lineal corrigiendo la endogeneidad de la variable “apertura” utilizando la variable instrumental propuesta, “recepción”. Esto en línea con lo propuesto por Angrist, Imbens & Rubin (1996). Si bien no es lo ideal modelar una regresión lineal con una variable dependiente binaria, esto es una de las recomendaciones por Angrist & Pischke (2009) en el libro “*Mostly Harmless Econometrics*”, debido a que otras aproximaciones no son

correctas al generar en ciertas etapas “regresiones prohibidas” (*forbidden regressions*), y la estimación utilizando 2SLS obtiene el “CACE” (o “LATE” como menciona en el libro).

Dado lo anterior, el modelo a estimar es el siguiente,

Ecuación 6: Especificación del modelo regresión lineal – simple de la segunda etapa, para evaluación del experimento, enfoque CACE.

$$X_i = B_0 + \sum_{j=\{Ec, Ec-txPrs, Prs, Amb\}} B_j * Condición_{i,j} * \widehat{Abre}_i + e_i$$

Ecuación 7: Especificación de primera etapa de corrección de endogeneidad mediante variable instrumental, modelo simple, enfoque CACE.

$$Condición_{i,j} * \widehat{Abre}_i = Condición_{i,j} * Recibe_i$$

En donde, se considera:

- X_i : variable binaria que toma valores 1 o 0, dependiendo si la persona i reportó sobre un paradero.
- $Condición_{i,j}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i fue asignada a la condición experimental j , las cuales son “Económico”, “Económico – texto Prosocial”, “Prosocial” o “Ambos”.
- $Abre_i$: variable binaria que toma valores 1 o 0, dependiendo si la persona i abre o no la notificación que recibió.
- $Recibe_i$: variable binaria que toma valores 1 o 0, dependiendo si la persona i recibe o no la notificación a la cual fue asignada a recibir.

En la Ecuación 7 se indica la primera etapa de la estimación del modelo, en donde la variable endógena, “Abre” (apertura), se ve estimada mediante OLS utilizando como variable instrumental “Recibe” (recepción). Estos resultados son luego empleados para la estimación de la segunda etapa, como se observa en la Ecuación 6, en donde se obtienen los resultados de la estimación que se muestran en la Tabla 66.

Tabla 66: Resultados modelo lineal - simple, para estimación del efecto CACE.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) | Efecto marginal |
|---------------------|------------------|-----------------------|----------------|--------------------|------------------------|
| “Económico” | 0,3588 | 0,016 | 22,46 | 0,000 | 35,88% |
| “Econ. – txt. Pr.” | 0,3305 | 0,017 | 19,84 | 0,000 | 33,05% |
| “Prosocial” | 0,1661 | 0,041 | 4,08 | 0,000 | 16,61% |
| “Ambos” | 0,2851 | 0,020 | 14,08 | 0,000 | 28,51% |
| Intercepto | 0,0067 | 0,002 | 2,71 | 0,007 | - |

Como la regresión es lineal, los estimadores corresponden a los porcentajes de los efectos marginales (última columna). Por otro lado, se observa que todos los coeficientes son positivos y significativos, por lo que la apertura de las notificaciones provocan que la probabilidad de emitir un reporte aumente. Por ejemplo, la notificación con incentivo “Económico” aumenta en un 35,88% la probabilidad, siendo el más alto, y la notificación “Prosocial” un 16,61% siendo la menos efectiva.

En la Tabla 67, se presenta los resultados de la comparación por pares para así observar diferencias entre apertura de las notificaciones.

Tabla 67: Resultados análisis comparación por pares entre notificaciones, del modelo simple en efecto CACE

| Condiciones experimentales | Contraste | Valor z | Pr(> z) |
|-----------------------------------|------------------|----------------|--------------------|
| “Económico” vs “Econ. – txt. Pr.” | 0,028 | 1,58 | 0,115 |
| “Económico” vs. “Prosocial” | 0,192 | 5,00 | 0,000 |
| “Económico” vs. “Ambos” | 0,073 | 3,64 | 0,000 |
| “Econ. – txt. Pr.” vs “Prosocial” | 0,164 | 4,26 | 0,000 |
| “Econ. – txt. Pr.” vs “Ambos” | 0,045 | 2,21 | 0,027 |
| “Prosocial” vs “Ambos” | -0,118 | -3,05 | 0,002 |

Se concluye que hay evidencia de que todas las notificaciones son significativas entre sí con una confianza del 95%, a excepción de la “Económico” con “Económico – texto Prosocial”. Lo que simboliza, preliminarmente, que no se puede concluir que el contenido de la notificación provoque diferencias en la probabilidad de emitir un reporte.

5.5.2.2 Análisis – modelo según tipo de usuario

En esta sección se procede a analizar si el efecto de la apertura de las notificaciones presenta heterogeneidad dependiendo del tipo de usuario, por lo que se procede a calcular el efecto “CACE” primero de manera teórica (ver Tabla 68 y Tabla 69), y luego realizando un modelo de regresión lineal como en el caso anterior, para así obtener la significancia estadística de las condiciones experimentales.

Tabla 68: Estadísticas para estimación del ITT - modelo tipo de usuario.

| Tipo de usuario | Condición experimental | Usuarios asignados | Reciben al menos una notificación | Reportan al menos una vez | ITT |
|---------------------|------------------------|--------------------|-----------------------------------|---------------------------|--------|
| “Reportador” | “Económico” | 274 | 119 (43,43%) | 46 (38,66%) | 28,60% |
| | “Econ. – txt. Pr.” | 274 | 122 (44,52%) | 41 (33,60%) | 23,54% |
| | “Prosocial” | 160 | 66 (41,25%) | 25 (37,88%) | 27,82% |
| | “Ambos” | 274 | 100 (36,49%) | 29 (29,00%) | 18,94% |
| | “Control” | 159 | - | 16 (10,06%) | - |
| “Reportador Tímido” | “Económico” | 1.679 | 694 (41,33%) | 104 (14,99%) | 13,56% |
| | “Econ. – txt. Pr.” | 1.679 | 662 (39,42%) | 104 (15,71%) | 14,28% |
| | “Prosocial” | 979 | 409 (41,77%) | 13 (3,18%) | 1,75% |
| | “Ambos” | 1.679 | 637 (37,94%) | 88 (13,91%) | 12,48% |
| | “Control” | 979 | - | 14 (1,43%) | - |
| “No Reportador” | “Económico” | 9.211 | 3.410 (37,02%) | 255 (7,48%) | 7,24% |
| | “Econ. – txt. Pr.” | 9.211 | 3.485 (37,83%) | 219 (6,28%) | 6,04% |
| | “Prosocial” | 5.373 | 1.964 (36,55%) | 26 (1,32%) | 1,08% |
| | “Ambos” | 9.211 | 3.497 (37,97%) | 148 (4,23%) | 3,99% |
| | “Control” | 5.373 | - | 13 (0,24%) | - |

Tabla 69: Estadísticas para estimación del CACE - modelo tipo de usuario.

| Tipo de usuario | Condición experimental | Reciben al menos una notificación | Abren al menos una notif. [ITT _D] | Reportan al menos una vez | CACE |
|---------------------|------------------------|-----------------------------------|---|---------------------------|--------|
| “Reportador” | “Económico” | 119 | 49 (41,18%) | 33 (67,35%) | 69,45% |
| | “Econ. – txt. Pr.” | 122 | 54 (44,26%) | 32 (59,26%) | 53,18% |
| | “Prosocial” | 66 | 24 (36,36%) | 16 (66,67%) | 76,51% |
| | “Ambos” | 100 | 40 (40,00%) | 22 (55,00%) | 47,35% |
| | “Control” | - | - | 16 (10,06%) | - |
| “Reportador Tímido” | “Económico” | 694 | 231 (33,29%) | 85 (36,80%) | 40,73% |
| | “Econ. – txt. Pr.” | 662 | 193 (29,15%) | 88 (45,60%) | 48,99% |
| | “Prosocial” | 409 | 56 (13,69%) | 7 (12,50%) | 12,78% |
| | “Ambos” | 637 | 199 (31,24%) | 64 (32,16%) | 39,95% |
| | “Control” | - | - | 14 (1,43%) | - |
| “No Reportador” | “Económico” | 3.410 | 769 (22,55%) | 230 (29,90%) | 32,10% |
| | “Econ. – txt. Pr.” | 3.485 | 767 (22,00%) | 190 (24,77%) | 27,45% |
| | “Prosocial” | 1.964 | 206 (10,49%) | 15 (7,28%) | 10,30% |
| | “Ambos” | 3.497 | 590 (16,87%) | 124 (21,02%) | 23,65% |
| | “Control” | - | - | 13 (0,24%) | - |

Luego, para obtener los resultados econométricos, se enuncia las especificaciones del modelo 2SLS a estimar en las siguientes ecuaciones.

Ecuación 8: Especificación del modelo regresión lineal – tipo de usuario de la segunda etapa, para evaluación del experimento, enfoque CACE.

$$X_i = \sum_{j=\{Ec, Ec-txPrs, Prs, Amb\}} \sum_{k=\{NoRep, RepTim, Rep\}} B_j * Condición_{i,j} * TipoUsuario_{i,k} * \widehat{Abre}_i + \sum_{l=\{NoRep, RepTim, Rep\}} B_l * Control_{i,l} + e_i$$

Ecuación 9: Especificación de primera etapa de corrección de endogeneidad mediante variable instrumental, modelo tipo de usuario, enfoque CACE.

$$Condición_{i,j} * TipoUsuario_{i,k} * \widehat{Abre}_i = Condición_{i,j} * TipoUsuario_{i,k} * Recibe_i$$

En donde, se considera:

- X_i : variable binaria que toma valores 1 o 0, dependiendo si la persona i reportó sobre un paradero.
- $Condición_{i,j}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i fue asignada a la condición experimental j, las cuales son “Económico”, “Económico – texto Prosocial”, “Prosocial” o “Ambos”.
- $TipoUsuario_{i,k}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i es usuario tipo “No Reportador”, “Reportador Tímido” o “Reportador”.
- $Control_{i,l}$: variable binaria que toma valores 1 o 0, dependiendo si la persona i fue asignada al grupo de control l, el cual puede ser “Grupo de control – No Reportador”, “Grupo de control – Reportador Tímido” y “Grupo de control – Reportador”.
- $Abre_i$: variable binaria que toma valores 1 o 0, dependiendo si la persona i abre o no la notificación que recibió.
- $Recibe_i$: variable binaria que toma valores 1 o 0, dependiendo si la persona i recibe o no la notificación a la cual fue asignada a recibir.

Notar que en la última sumatoria de la Ecuación 8, sólo se deben considerar dos clasificaciones de tipo de usuario, ya que la tercera corresponde al caso base de la ecuación. Los resultados de esta regresión lineal, se adjuntan en la Tabla 70, en donde el caso base corresponde a los usuarios “No Reportadores”. El resto de las tablas (con caso base “Reportadores Tímidos” y “Reportadores”) se adjuntan en Anexos, Apéndice N: Resultados de regresión lineal 2SLS – para estimación CACE modelo tipo de usuario, en donde se desprende que todas las condiciones experimentales en la categoría de usuario “Reportador” son significativas respecto al grupo de control de los usuarios “Reportador”, lo cual recíproco para la categoría de “Reportadores Tímidos”.

Tabla 70: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, con grupo base control del tipo de usuario “No Reportador”.

| Coeficiente | | Estimador | Error estándar | Valor Z | Pr(> z) | Efecto marginal |
|---------------------|--------------------|-----------|----------------|---------|----------|-----------------|
| “Reportador” | “Económico” | 0,9387 | 0,044 | 21,17 | 0,000 | 93,87% |
| | “Econ. – txt. Pr.” | 0,7592 | 0,041 | 18,64 | 0,000 | 75,92% |
| | “Prosocial” | 1,0416 | 0,067 | 15,45 | 0,000 | 104,16% |
| | “Ambos” | 0,7250 | 0,050 | 14,56 | 0,000 | 72,50% |
| | “Control” | 0,1006 | 0,016 | 6,37 | 0,000 | 10,06% |
| “Reportador Tímido” | “Económico” | 0,4502 | 0,023 | 19,82 | 0,000 | 45,02% |
| | “Econ. – txt. Pr.” | 0,5389 | 0,027 | 20,30 | 0,000 | 53,89% |
| | “Prosocial” | 0,2321 | 0,072 | 3,23 | 0,001 | 23,21% |
| | “Ambos” | 0,4422 | 0,025 | 17,51 | 0,000 | 44,22% |
| | “Control” | 0,0143 | 0,006 | 2,25 | 0,025 | 1,43% |
| “No Reportador” | “Económico” | 0,3315 | 0,015 | 21,93 | 0,000 | 33,15% |
| | “Econ. – txt. Pr.” | 0,2855 | 0,015 | 18,63 | 0,000 | 28,55% |
| | “Prosocial” | 0,1262 | 0,043 | 2,95 | 0,003 | 12,62% |
| | “Ambos” | 0,2508 | 0,020 | 12,57 | 0,000 | 25,08% |

En la Tabla 70, se observa que todas los coeficientes asociados a notificaciones son positivos y estadísticamente significativos, por lo que la apertura de una notificación aumenta la probabilidad de emitir un reporte respecto al grupo de control de la categoría “No Reportador”. Esta magnitud depende del tipo de usuario, por lo cual se procede a realizar un cambio de base en la Tabla 71, para así aislar el efecto CACE, y donde se compara respecto al efecto estimado teóricamente (calculado utilizando la fórmula de la Ecuación 5: Identidad para cálculo de efecto CACE).

Tabla 71: Resultados de efecto CACE a partir de modelo – tipo de usuario, y contraste con el teórico

| Condición experimental | Reportadores | | Reportadores Tímidos | | No Reportadores | |
|------------------------|--------------|---------|----------------------|---------|-----------------|---------|
| | 2SLS | Teórico | 2SLS | Teórico | 2SLS | Teórico |
| “Económico” | 83,81% | 69,45% | 43,59% | 40,73% | 33,15% | 32,10% |
| “Econ. – txt. Pr.” | 65,86% | 53,18% | 52,46% | 48,99% | 28,55% | 27,45% |
| “Prosocial” | 94,10% | 76,51% | 21,78% | 12,78% | 12,62% | 10,30% |
| “Ambos” | 62,44% | 47,35% | 42,79% | 39,95% | 25,08% | 23,65% |
| “Control” | - | - | - | - | - | - |

Al analizar primero el contraste entre los efectos estimados mediante 2SLS y los teóricos (Tabla 69), se tiene que en la categoría “No Reportador” es mínima la diferencia, la cual aumenta de manera importante en las otras categorías. Esto se debe a las limitaciones del modelo 2SLS al utilizar una variable dependiente binaria con baja cantidad de usuarios.

Se desprende que el efecto de la apertura (y se asume, lectura) de la notificación provoca un efecto positivo en la probabilidad de emitir un reporte, el cual aumenta de magnitud a mayor uso asociado a la clasificación del tipo de usuario en su historial de utilización de la herramienta de envío de reportes en la aplicación. Al igual que en el enfoque “ITT”, el efecto “Prosocial” es el más alto en los “Reportadores” y el más bajo en los “Reportadores Tímidos” y en los “No Reportadores”.

A continuación se realiza un análisis de comparación por pares para identificar diferencias estadísticas significativas entre los efectos de la apertura de las notificaciones (ver Tabla 72).

Tabla 72: Resultados análisis comparación por pares entre notificaciones, del modelo tipo de usuario en efecto CACE

| Tipo de usuario | Condiciones experimentales | Contraste | Valor z | Pr(> z) |
|---------------------|-----------------------------------|-----------|---------|----------|
| “Reportador” | “Económico” vs “Econ. – txt. Pr.” | 0,179 | 2,98 | 0,003 |
| | “Económico” vs. “Prosocial” | -0,102 | -1,27 | 0,205 |
| | “Económico” vs. “Ambos” | 0,213 | 3,22 | 0,001 |
| | “Econ. – txt. Pr.” vs “Prosocial” | -0,281 | -3,58 | 0,000 |
| | “Econ. – txt. Pr.” vs “Ambos” | 0,035 | 0,54 | 0,587 |
| | “Prosocial” vs “Ambos” | 0,316 | 3,78 | 0,000 |
| “Reportador Tímido” | “Económico” vs “Econ. – txt. Pr.” | -0,087 | -2,51 | 0,012 |
| | “Económico” vs. “Prosocial” | 0,229 | 3,01 | 0,003 |
| | “Económico” vs. “Ambos” | 0,008 | 0,25 | 0,802 |
| | “Econ. – txt. Pr.” vs “Prosocial” | 0,317 | 4,10 | 0,000 |
| | “Econ. – txt. Pr.” vs “Ambos” | 0,096 | 2,63 | 0,009 |
| | “Prosocial” vs “Ambos” | -0,220 | -2,87 | 0,004 |
| “No Reportador” | “Económico” vs “Econ. – txt. Pr.” | 0,046 | 2,16 | 0,031 |
| | “Económico” vs. “Prosocial” | 0,218 | 4,61 | 0,000 |
| | “Económico” vs. “Ambos” | 0,084 | 3,34 | 0,001 |
| | “Econ. – txt. Pr.” vs “Prosocial” | 0,172 | 3,64 | 0,000 |
| | “Econ. – txt. Pr.” vs “Ambos” | 0,038 | 1,51 | 0,132 |
| | “Prosocial” vs “Ambos” | -0,134 | -2,78 | 0,005 |

Se observa que,

- En los usuarios “Reportadores” todas las notificaciones entre sí tienen diferencias significativas, excepto en los casos de las notificaciones “Económico” y “Prosocial”, y “Económico – texto Prosocial” y “Ambos”.
- Para la categoría de usuarios “Reportadores Tímidos”, sólo las notificaciones “Económico” y “Ambos” no presentan diferencia estadística significativa entre sí.
- En los tipos de usuarios “No Reportadores” todas presentan diferencias significativas entre sí exceptuando las notificaciones “Económico – texto Prosocial” vs “Ambos”.

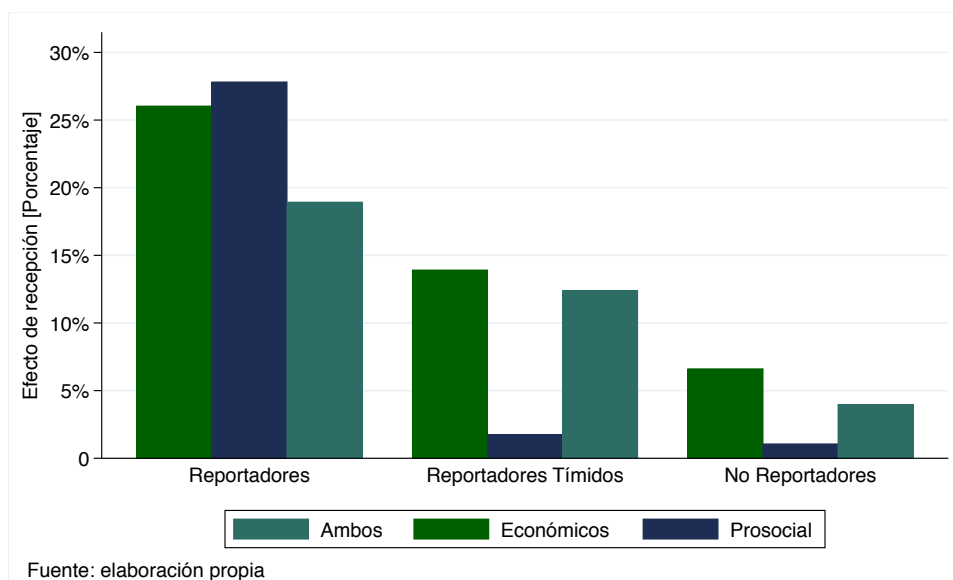
5.6 Conclusiones del experimento de campo

Ya sea observando los resultados de los modelos según recepción o apertura, los resultados más reveladores se encuentran al considerar información según la clasificación según tipo de usuario, ya que evidencia una fuerte heterogeneidad del tratamiento, por lo que se procede a enunciar conclusiones en base a éstos modelos.

5.6.1 Conclusiones según recepción

Los resultados obtenidos utilizando el enfoque de recepción son los más “limpios”, ya que son en base a sólo a la recepción de alguna notificación, independiente si la persona se autoseleccionó para recibir el tratamiento intensivo y así generando un sesgo de selección. En la Ilustración 39 se muestra resumidamente los efectos obtenidos en la estimación según tipo de usuario. Posteriormente, se menciona los principales efectos encontrados en cada una de la de las categorías de usuarios.

Ilustración 39: Resultados resumidos del efecto ITT - Recepción en la probabilidad de emitir reporte.



En los usuarios “Reportadores” si bien la notificación “Prosocial” es la que tiene un efecto más alto (27,81%), no presenta diferencias estadísticas significativas respecto al resto de las notificaciones. Al aumentar en una cifra similar la probabilidad de emitir un reporte como las notificaciones con incentivos económicos (“Económicos” (26,03%) y “Ambos” (18,93%)), se considera como la alternativa más costo-eficiente, ya que no presenta un coste monetario para la aplicación.

En los usuarios “Tímidos” la situación es diferente, ya que la notificación “Prosocial” si bien tiene un efecto de 1,75%, éste es bastante inferior al resto de las notificaciones que involucran un

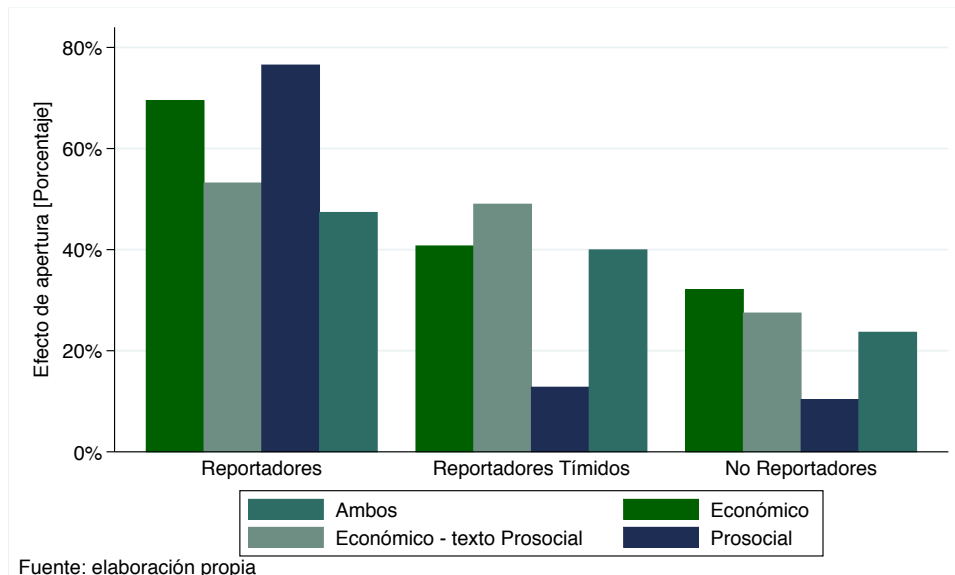
incentivo económico, los cuales son 13,91% y 12,39% para las notificaciones “Económicos” y “Ambos”, respectivamente. Notar que no hay diferencias estadísticas entre éstas, por lo que se concluye que las personas reaccionan de manera más efectiva a los incentivos que contienen una componente económica.

En la categoría de usuarios menos “activos”, los “No Reportadores”, las notificaciones tienen un efecto menor a lo observado en las previas categorías, en especial aquellas que involucran un incentivo económico. Al analizar el efecto entre las diferentes notificaciones, las “Económicos” otorgan un aumento en la probabilidad de reportar de 6,61%, el cual es cerca de seis veces más efectivo que el “Prosocial” y el doble que el obtenido en la notificación “Ambos”. De igual forma que en la categoría de usuarios “Tímidos”, los usuarios responden en mayor medida a los incentivos económicos.

5.6.2 Conclusiones según apertura

Las conclusiones de los resultados del enfoque según apertura, si bien son válidos, son acotados a la población que finalmente decidió abrir las notificaciones. Es por esto que el “CACE” también se le denomina como “Local Average Treatment Effect” (Efecto promedio del tratamiento local, en español) al ser un efecto observado en ciertos tipos de usuarios y no en toda la muestra. De igual forma que en la sección previa, se concluye progresivamente en base a la clasificación de tipo de usuario. Los efectos encontrados se grafican en la Ilustración 40.

Ilustración 40: Resultados teóricos resumidos del efecto CACE - Apertura en la probabilidad de emitir reporte.



En los usuarios “Reportadores” se tiene que la apertura de una notificación “Prosocial” es la que tiene un efecto más alto de 76,51%, el cual no tiene diferencias estadísticas respecto al incentivo “Económico” de 69,45%. En base a esta información, e igual en el caso del enfoque según apertura, la opción “Prosocial” sería la más eficiente al no considerar un coste monetario a *Transapp*. Por

otro lado, al comparar las notificaciones “Económico” y “Económico – texto Prosocial” que tenían la misma notificación, al ver el efecto de la lectura de su mensaje, el cual difería, se tienen efectos estadísticamente diferentes, en donde agregar una componente prosocial disminuyó la probabilidad de emitir un reporte de 69,45% a 53,18%. Este efecto es incluso menor que la notificación “Prosocial” disminuyendo de un 76,51% a un 53,18%. Dado lo anterior, se recomienda no combinar incentivos en este tipo de usuarios ya que tienen un efecto menor que en solitario.

Para la categoría de usuarios “Tímidos”, se tiene que la notificación “Prosocial” es la con menor efecto, cercano a la mitad de las otras que involucran un incentivo económico. Destaca en este escenario, en contraste al anterior, que la notificación “Económico – texto Prosocial” tiene un efecto más alto que el de “Económico” (48,99% versus 40,73%, respectivamente), lo que significa que la componente Prosocial del mensaje aumentó la motivación de las personas en reportar.

Finalmente, en los usuarios “No Reportadores”, la notificación “Prosocial” nuevamente es la menos efectiva y la “Económico” es la con el efecto más alto de la lectura del mensaje. Respecto a la diferencia de las notificaciones “Económico” y “Económico – texto Prosocial”, esta es de un 32,10% a un 27,45% cuya diferencia es significativa, por lo que la inclusión de una componente prosocial disminuye la motivación de las personas en emitir un reporte.

5.6.3 Conclusiones teóricas

En general se concluye que para quienes realizan regularmente la actividad de reportes, un incentivo prosocial es el óptimo, a diferencia de quienes no lo hacen regularmente (o nunca) a quienes hay que motivar utilizando incentivos económicos, a pesar de que todos los incentivos recibidos son positivos para todos los tipos de usuario.

La teoría indica la existencia de dos tipos de motivaciones en la realización de acciones/comportamientos: la intrínseca y la extrínseca. La primera indica la motivación en la cual el beneficio es propio de la acción (White, 1959) y la segunda refiere a cuando el beneficio es externo y diferente a la actividad (Schwartz, Milfont, & Hilton, 2019).

En los usuarios que regularmente reportan, se supone que es debido a una acción altruista y/o por su participación de la *gamificación*, por lo que los usuarios responden a las dos motivaciones en conjunto. Al incentivar a los usuarios ya sea mediante un incentivo prosocial (motivación intrínseca) o incentivo económico (motivación extrínseca) la proporción de reportadores aumenta en gran magnitud. Lamentablemente la muestra es pequeña por lo que no es posible identificar diferencias significativas entre ellas. Lo que sí es interesante es que éstos usuarios tienen una mayor tasa de lectura que el resto de las clasificaciones de usuarios, y que la tasa de lectura es similar entre notificaciones, a diferencia de las otras. Por ejemplo, la tasa de apertura está comprendida entre 36% y 45% (dependiendo de la notificación), la cual se reduce a 10% y 23% en los usuarios “No Reportadores”, esto puede explicarse debido a un mayor interés en la actividad prosocial (asociado a mayor motivación intrínseca) y a mayor familiaridad con la aplicación (debido a mayor uso). En

conclusión, en esta categoría los usuarios responden a ambos incentivos de manera similar (lo cual ya es diferenciador al resto de las clasificaciones), no obstante, destaca el mayor interés por acciones prosociales el cual se manifiesta primero en mayor apertura de la notificación (y menor diferencia respecto a las económicas), y segundo en un fuerte incremento en la proporción de usuarios reportadores que reciben esta notificación.

Por el otro lado, en los usuarios que no reportan (o lo hacen infrecuentemente) se presume que no realizan la acción por desconocimiento y/o falta de interés al no observar una utilidad (en caso de que sí reportaran, sería por su motivación intrínseca y/o el incentivo de *gamificación*, como se indica en el párrafo anterior). Luego, al destacar un incentivo prosocial, si bien aumenta levemente la motivación intrínseca (ya que aumentan los reportes realizados) ésta no es de gran magnitud. En cambio, al incentivar mediante un incentivo económico externo, la proporción de personas reportadoras aumenta más fuertemente ya que éste provoca un aumento de su motivación extrínseca por lo que realizan la acción por el deseo de obtener el premio de la lotería. Notar que en ambos escenarios se les explica a los usuarios cómo reportar de manera idéntica, por lo que la diferencia entre ellos se debe netamente por el tipo de incentivo (aunque esto solo se aprecia en el contenido de la notificación, es decir, en el efecto “CACE”). De manera particular, sólo en la categoría de “No Reportadores” se evidencia una diferencia significativa en la proporción de reportadores entre la notificación “Económico” y “Ambos”, por lo que se concluye que un incentivo económico combinado es perjudicial respecto a uno plenamente económico debido a que el primero tiene un efecto positivo de menor magnitud que el segundo. Hay indicios de lo anterior en las otras condiciones, no obstante no hay evidencia estadística para afirmar aquello (lo cual se debe a una baja muestra).

Al considerar el efecto de la apertura de la notificación, sólo es comparable el efecto entre las condiciones “Económico” y “Económico-texto prosocial”. Dado que los usuarios deciden abrir la notificación es porque el título y subtítulo de la notificación les causa interés, luego, se concluye que son usuarios que se ven incitados debido al premio de la lotería, el incentivo externo. Por lo cual, al incluir una componente prosocial en el contenido del texto (sumada a la económica) es esperable una disminución en la proporción de usuarios reportadores ya que su interés radica en una motivación extrínseca y no en una intrínseca. Esta situación se cumple en los usuarios “Reportadores” y “No reportadores”, pero no así en los “Reportadores tímidos”.

5.7 Discusión y futuros estudios

Si bien el estudio fue realizado exitosamente en este trabajo de título, hubo aprendizajes basados en limitaciones del experimento que se describen a continuación para que investigadores/as lo puedan considerar en sus futuros trabajos.

Durante la realización del experimento, en caso de no tener la posibilidad de aislar las acciones de las personas de manera efectiva, se recomienda no enviar otras notificaciones a los usuarios para así poder aislar diferencias en comportamientos en los diferentes grupos.

Considerar en los grupos experimentales un grupo de placebo, que permita identificar si el efecto en las diferencias de comportamiento de los usuarios es debido a la recepción de la notificación o al contenido de ésta. En este trabajo de título no se considera debido a que disminuiría la muestra de las otras condiciones experimentales provocando que fuese menos probable observar efectos significativos.

Procurar que la persona reciba la notificación cuando esté usando la aplicación (o que sea lo más próximo posible), para así poder observar la reacción del usuario dado que está en un entorno en el cual efectivamente se trata de una intervención en la acción que se desea cambiar.

Aumentar el tamaño muestral (en relación con este experimento) para así tener más probabilidades de evidenciar estadísticamente un efecto pequeño en la intervención. De igual manera, también permitiría crear nuevas condiciones experimentales.

Sería interesante estudiar cuál es la intensidad óptima del tratamiento y ver si este varía de manera significativa entre más o menos notificaciones recibe el usuario. Considerando crítico que el envío de notificaciones es una actividad que en exceso es *molesta* a los usuarios dado que es la razón por la cual el 71% de las aplicaciones son desinstaladas (Appiterate, 2013). Notar que en *Transapp* la cantidad de desinstalaciones no aumentó durante el envío de las notificaciones del experimento (lo cual era una interrogante importante para el equipo desarrollador).

Analizar si los tratamientos tienen un efecto breve o prolongado en el tiempo (de mayor duración que en el experimento) en el comportamiento de las personas en la aplicación, lo que no pudo ser realizado en esta investigación por la crisis social ocurrida en Chile en 2019, ya que las condiciones de transporte y estabilidad social cambiaron drásticamente el día siguiente al terminar el experimento.

Capítulo 6

Conclusiones generales

En este trabajo de título se inicia un primer paso para entender el comportamiento de usuarios de una aplicación para teléfonos inteligentes relacionada al transporte público.

En primer lugar se realiza un análisis de datos que permite visualizar la actividad de los usuarios en las principales funciones de la aplicación, como en las consultas de los paraderos, los reportes de infraestructura y los viajes realizados en bus. Se indentifican estadísticas semanales, variaciones locales según intervalo horario y día de la semana, se realiza un modelo que explica variaciones en la evaluación entregada por los usuarios de los viajes en la aplicación y se confirma mediante el trabajo de datos la presencia de la Regla de Nielsen en *Transapp*. Este trabajo con los usuarios, fue también parte del análisis de la sobrevivencia de los usuarios, en donde se muestra que los usuarios que utilizan más herramientas en su primera semana tienen menor probabilidad de volverse inactivos.

En segundo lugar, se realiza un experimento de campo que ofrece incentivos económicos y/o prosociales a los usuarios para aumentar la participación en la realización de reportes sobre los paraderos de la ciudad. En este se evidencia una fuerte heterogeneidad del tratamiento, en donde se recomienda utilizar incentivos prosociales por ser coste-efectivos en los usuarios más activos (en términos de reportes realizados previamente) a diferencia de los usuarios menos activos, a quienes se recomienda motivar mediante incentivos monetarios.

Por otro lado, este trabajo de título también genera aprendizajes para el equipo desarrollador de la aplicación, los cuales les ha permitido potenciar canales informativos en su cadena de valor de la aplicación. Por ejemplo, se descubre que el texto de la notificación push debe ser atractivo para captar la atención del usuario y que éste la abra, también existe evidencia de que los usuarios sí leen el mensaje de las notificaciones, y que un envío regular de notificaciones no es nocivo en términos de impactar negativamente y aumentar la cantidad de desinstalaciones en el corto plazo.

Finalmente, tener una mayor comprensión de los usuarios es de suma importancia, ya que permite desarrollar estrategias que estimulen al usuario a seguir utilizando (y en mayor medida) la aplicación lo cual es crítico en términos de ventajas comparativas respecto a la competencia, en un contexto muy competitivo donde cambiarse de aplicación sólo requiere un impulso de segundos. De igual manera, también permite desarrollar nuevas herramientas que potencian la capacidad que tienen los usuarios de ser agentes activos en la mejora del sistema de transporte, a través de demandas sobre un transporte público de mejor calidad.

Bibliografía

- Angrist, J., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton: Princeton University Press.
- Angrist, J., Imbens, G., & Rubin, D. (1996). Identification of Causal Effects Using Instrumental Variables. *Journal of the American Statistical Association* Vol. 91 No.434, 444-455.
- Antoine, A. (2018). *Caso práctico de estudio de incorporación de técnicas de gamificación en aplicación móvil, TranSapp*. Obtenido de Repositorio Académico de la Universidad de Chile: <http://repositorio.uchile.cl/handle/2250/152918>
- Appiterate. (23 de Diciembre de 2013). *Top 7 reasons why mobile users uninstall apps*. Obtenido de PRLOG Press Release Distribution: <https://www.prlog.org/12254771-top-7-reasons-why-mobile-users-uninstall-apps.html>
- Baym, N. (1993). Interpreting Soap Operas and Creating Community: Inside a Computer-Mediated Fan Culture. *Journal of Folklore Research*, 143 - 176.
- Bovornkeeratiroj, P., Nakorn, K. N., & Rojviboonchai, K. (2015). Challenges in implementing crowdsourcing on automatic real-time transit tracking system. *2015 15th International Symposium on Communications and Information Technologies (ISCIT)*, 5-8.
- Brakewood, C., Macfarlane, G., & Watkins, K. (2015). The impact of real-time information on bus ridership in New York City. *Transportation Research Part C: Emerging Technologies*, Vol. 53, 59-75.
- Brazil, W., & Caulfield, B. (2013). Does green make a difference: The potential role of smartphone technology in transport behaviour. *Transportation Research Part C: Emerging Technologies*, 93-101.
- Bronner, F., & Hoog, R. d. (2011). Vacationers and eWOM: Who Posts, and Why, Where, and What? *Journal of Travel Research* Vol. 50, 15-26.
- Brújula. (2017). *Estudio satisfacción de empresas operadoras*. Obtenido de DPTM: <http://www.dtpm.cl/descargas/estudios/Informe%20Final%20Satisfacci%C3%B3n%20Empresas%20Operadoras%202017.pdf>
- Burke, M., Marlow, C., & Lento, T. (2009). Feed Me: Motivating Newcomer Contribution in Social Network Sites. *CHI '09 Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, (págs. 945 - 954). Boston, MA.
- Cleves, M., Gould, W., & Marchenko, Y. (2008). *An Introduction to Survival Analysis Using Stata*. Stata press.
- CNN Chile. (12 de Marzo de 2019). *CNN Chile*. Obtenido de Nueva flota de buses y paraderos inteligentes: Red, el sistema de transporte que reemplazará al Transantiago: https://www.cnnchile.com/pais/red-sistema-transporte-publico-transantiago_20190312/
- Correa, P. (8 de Febrero de 2017). *T13*. Obtenido de Un día en el recorrido más largo del Transantiago: <https://www.t13.cl/noticia/nacional/un-dia-en-el-recorrido-mas-largo-del-transantiago>

- De Oña, J., De Oña, R., Eboli, L., & Mazzullaba, G. (2013). Perceived service quality in bus transit service: A structural equation approach. *Transport Policy Vol. 29*, 219 - 226.
- Dickinson, J., Cherrett, T., Hibbert, J., Winstanley, C., Duncan, S., Davies, N., . . . Speed, C. (2015). Fundamental challenges in designing a collaborative travel app. *Transport Policy Vol. 44*, 28-36.
- Dixit, S., Badgaiyan, A. J., & Khare, A. (2019). An integrated model for predicting consumer's intention to write onlinereviews. *Journal of Retailing and Consumer Services 46*, 112-120.
- DTPM. (2019). *Informe de Gestión 2018*. Obtenido de http://www.dtpm.cl/descargas/memoria/Informe_Gestion_2018_DTPM.pdf
- Eboli, L., & Mazzullaba, G. (2007). Service Quality Attributes Affecting Customer Satisfaction for Bus Transit. *Journal of Public Transportation, Vol. 10*, 21 - 34.
- EUR-LEX. (7 de Julio de 2010). *El acceso al Drecho de la Unión Europea*. Obtenido de Directive 2010/40/EU of the European Parliament and of The Council of 7 July 2010: <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2010:207:0001:0013:EN:PDF>
- Faber, A., & Matthes, F. (2016). Crowdsourcing and Crowdinnoation. *Digital Mobility Platforms and Ecosystems*, 36-47.
- Ferris, B., Watkins, K., & Borning, A. (2010). OneBusAway: results from providing real-time arrival information for public transit. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 1807-1816.
- Filippi, F., Fusco, G., & Nanni, U. (2013). User empowerment and advanced public transport solutions. *Procedia - Social and Behavioral Sciences, Vol. 87*, 3-17.
- Google. (2019). *Introduction to Push Notifications*. Obtenido de Progressive Web Apps Training: <https://developers.google.com/web/ilt/pwa/introduction-to-push-notifications>
- Ling, K., Beenen, G., Ludford, P., Wang, X., Chang, K., Li, X., . . . Kraut, R. (2006). Using Social Psychology to Motivate Contributions to Online Communities. *Journal of Computer-Mediated Communication (10) 4*, Artículo 10.
- Litvin, S., Goldsmith, R., & Pan, B. (2008). A Literature Review of Word of Mouth and Electronic Word of Mouth: Implications for Consumer Behavior. *Tourism Management Vol. 29, Issue 3*, 458 - 468.
- Lu, H., Burge, P., Heywood, C., Sheldon, R., Lee, P., Barber, K., & Phillips, A. (2018). The impact of real-time information on passenger's value of bus waiting time. *Transportation Research Procedia, Vol. 31*, 18-34.
- Malinen, S. (2015). Understanding User Participation in Online Communities: A Systematic Literature Review of Empirical Studies. *Computres in Human Behavior Vol. 46*, 228 - 238.
- Misra, A., Gooze, A., Watkins, K., Asad, M., & Le Dantec, C. (2014). Crowdsourcing and Its Application to Transportation Data Collection and Management. *Transportation Research Record 2414(1)*, 1-8.
- Nandan, N., Pursche, A., & Zhe, X. (2014). Challenges in crowdsourcing real-time information for public transportation. *15th International Conference on Mobile Data Management*, 67-72.
- Nielsen, J. (8 de Octubre de 2006). *The 90-9-1 Rule for Participation Inequality in Social Media and Online Communities*. Obtenido de NN/g Nielsen Norman Group: <https://www.nngroup.com/articles/participation-inequality/>

- Nonnecke, B., Andrews, D., Preece, J., & Voutour, R. (2004). Online Lurkers Tell Why. *AMCIS 2004 Proceedings*, 321.
- OECD. (13 de Mayo de 2019). *An Introduction to Online Platforms and Their Role in the Digital Transformation*. Obtenido de OECDiLibrary: https://www.oecd-ilibrary.org/science-and-technology/an-introduction-to-online-platforms-and-their-role-in-the-digital-transformation_19e6a0f0-en
- Park, N., Sook Oh, H., & Kang, N. (2015). Effects of ego involvement and social norms on individuals' uploading intention on Wikipedia: A comparative study between the United States and South Korea. *Journal of the Association for Information Science and Technology Vol. 66*, 1494 - 1506.
- Pfeifer, P., Haskins, M., & Conroy, R. (2005). Customer Lifetime Value, Customer Profitability, and the Treatment of Acquisition Spending. *Journal of Managerial Issues Vol.17, No.1*, 11-25.
- Rashid, A. M., Ling, K., Tassone, R. D., Resnick, P., Kraut, R., & Riedl, J. (2006). Motivating Participation by Displaying the Value of Contribution. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 955-958.
- Rosset, S., Neumann, E., Eick, U., Vatnik, N., & Idan, Y. (2002). Customer Lifetime Value Modeling and Its Use for Customer Retention Planning. *Proceedings of the eighth ACM SIGKDD international conference on Knowledge discovery and data mining*, 332-340.
- Schwartz, D., Milfont, T., & Hilton, D. (2019). The interplay between intrinsic motivation, financial incentives and nudges in sustainable consumption. En K. Gangl, & E. Kirchler, *A research agenda for economic psychology* (pág. 88). Edward Elgar Pub.
- Serra-Cantalops, A., & Salvi, F. (2014). New Consumer Behavior: A Review of Research on eWOM and Hotels. *International Journal of Hospitality Management Vol. 36*, 41 - 51.
- Singh, B., & Gupta, A. (2015). Recent trends in intelligent transportation systems: a review. *Journal of Transporte Literature, Vol. 9(2)*, 30-34.
- Suhonen, E., Lampinen, A., Cheshire, C., & Antin, J. (2010). Everyday Favors: a Case Study of a Local Online Gift Exchange System. *Group'10 Proceedings of the 16th ACM International Conference on Supporting Group Work*, 11-20.
- Sun, N., Pei-Luen Rau, P., & Ma, L. (2014). Understanding lurkers in online communities: A literature review. *Computers in Human Behavior Vol. 38*, 110 - 117.
- Sussman, J. (2005). *Perspectives on Intelligent Transportation Systems (ITS)*. Nueva York: Springer-Verlag.
- Swartz, A. (4 de Septiembre de 2006). *Who writes Wikipedia?* Obtenido de Raw Thought by Aaron Swartz: <http://www.aaronsw.com/weblog/howwriteswikipedia>
- Tang, L., & Thakuriah, P. (2012). Ridership effects of real-time bus information system: A case study in the City of Chicago. *Transportation Research Part C: Emerging Technologies, Vol. 22*, 146-161.
- Tarjeta bip! (2019). *¿Qué es tarjeta bip!?* Obtenido de Red bip!: <http://www.tarjetabip.cl/que-es-tarjeta-bip.php>
- Transapp. (Octubre de 2018). *¿Qué es Transapp?* Obtenido de Transapp: www.transapp.cl

- Universidad de Chile. (19 de Agosto de 2016). *TranSapp, la aplicación que busca mejorar los viajes en Transantiago*. Obtenido de Universidad de Chile: <http://www.uchile.cl/noticias/125223/transapp-la-aplicacion-que-busca-mejorar-los-viajes-en-transantiago>
- Vianna, F., Peinado, J., & Graeml, A. (2017). As motivações que levam usuários a participar de plataformas de crowdsourcing online. *XX Simpósio de Administração da produção, logística e operações internacionais*.
- Weitzenkorn, B. (14 de Junio de 2013). *Google to Buy Social Mapping Startup Waze*. Obtenido de NBC News: http://www.nbcnews.com/id/52207395/ns/business-small_business/t/google-buy-social-mapping-startup-waze/#.XYkdIZMzbkJ
- White, R. (1959). Motivation reconsidered: The concept of competence. *Psychological Review*, *66*(5), 297-333.

Anexos

Para facilidad de la lectura de este documento, al término de cada apéndice hay un hipervínculo de retorno al informe en donde se requiere este anexo.

1. Apéndice A: Estadísticas de consultas de paradero por semana

Tabla 73: Estadísticas descriptivas de las consultas de paraderos mensuales.

| Semana | Consultas realizadas | Consultas únicas | Usuarios activos | Usuarios que consultan | Proporción usuarios que consultan | Promedio consultas únicas por ID que consulta |
|--------|----------------------|------------------|------------------|------------------------|-----------------------------------|---|
| 1 | 544.974 | 46.554 | 7.200 | 7.200 | 100,00% | 6,47 |
| 2 | 875.281 | 72.780 | 9.141 | 8.391 | 91,80% | 8,67 |
| 3 | 837.529 | 68.662 | 9.709 | 8.233 | 84,80% | 8,34 |
| 4 | 954.335 | 75.205 | 10.325 | 8.573 | 83,03% | 8,77 |
| 5 | 830.875 | 66.262 | 10.695 | 8.547 | 79,92% | 7,75 |
| 6 | 1.076.980 | 85.430 | 12.030 | 9.946 | 82,68% | 8,59 |
| 7 | 1.011.687 | 90.493 | 13.233 | 10.967 | 82,88% | 8,25 |
| 8 | 618.124 | 56.109 | 12.955 | 8.544 | 65,95% | 6,57 |
| 9 | 1.019.754 | 84.855 | 13.936 | 10.768 | 77,27% | 7,88 |
| 10 | 1.164.770 | 97.049 | 14.833 | 11.778 | 79,40% | 8,24 |
| 11 | 1.173.821 | 98.567 | 15.386 | 12.041 | 78,26% | 8,19 |
| 12 | 1.012.299 | 93.576 | 15.966 | 12.368 | 77,46% | 7,57 |
| 13 | 1.222.390 | 106.405 | 16.559 | 12.939 | 78,14% | 8,22 |
| 14 | 753.573 | 70.761 | 16.656 | 11.829 | 71,02% | 5,98 |
| 15 | 1.253.751 | 114.302 | 18.011 | 14.076 | 78,15% | 8,12 |
| 16 | 1.339.580 | 118.246 | 18.946 | 14.684 | 77,50% | 8,05 |
| 17 | 1.341.971 | 120.771 | 19.538 | 14.879 | 76,15% | 8,12 |
| 18 | 1.060.361 | 101.516 | 19.853 | 14.414 | 72,60% | 7,04 |
| 19 | 1.286.501 | 119.109 | 20.648 | 15.448 | 74,82% | 7,71 |
| 20 | 1.329.809 | 123.117 | 21.221 | 15.717 | 74,06% | 7,83 |
| 21 | 1.280.239 | 123.852 | 22.025 | 16.330 | 74,14% | 7,58 |
| 22 | 1.110.723 | 107.038 | 22.340 | 15.622 | 69,93% | 6,85 |
| 23 | 951.782 | 93.689 | 22.552 | 14.849 | 65,84% | 6,31 |
| 24 | 1.226.524 | 128.081 | 23.589 | 16.563 | 70,21% | 7,73 |

| | | | | | | |
|----|-----------|---------|--------|--------|---------|------|
| 25 | 1.167.001 | 128.946 | 24.294 | 16.908 | 69,60% | 7,63 |
| 26 | 1.149.780 | 126.581 | 24.881 | 16.875 | 67,82% | 7,50 |
| 27 | 928.506 | 101.129 | 24.990 | 15.927 | 63,73% | 6,35 |
| 28 | 974.786 | 117.641 | 25.489 | 16.136 | 63,31% | 7,29 |
| 29 | 967.090 | 115.946 | 25.638 | 15.726 | 61,34% | 7,37 |
| 30 | 1.069.202 | 119.491 | 26.016 | 16.005 | 61,52% | 7,47 |
| 31 | 923.818 | 106.503 | 26.474 | 16.735 | 63,21% | 6,36 |
| 32 | 1.245.956 | 151.008 | 28.116 | 19.935 | 70,90% | 7,58 |
| 33 | 1.415.780 | 167.302 | 29.754 | 21.887 | 73,56% | 7,64 |
| 34 | 1.426.640 | 173.521 | 30.720 | 22.665 | 73,78% | 7,66 |
| 35 | 1.217.509 | 162.568 | 31.014 | 22.364 | 72,11% | 7,27 |
| 36 | 1.504.806 | 184.080 | 32.080 | 24.041 | 74,94% | 7,66 |
| 37 | 1.444.440 | 178.583 | 32.622 | 23.884 | 73,21% | 7,48 |
| 38 | 936.928 | 152.396 | 32.735 | 23.242 | 71,00% | 6,56 |
| 39 | 742.171 | 163.223 | 33.104 | 24.015 | 72,54% | 6,80 |
| 40 | 531.039 | 121.615 | 32.849 | 22.502 | 68,50% | 5,40 |
| 41 | 691.771 | 164.101 | 33.561 | 24.373 | 72,62% | 6,73 |
| 42 | 668.202 | 162.185 | 33.566 | 24.189 | 72,06% | 6,70 |
| 43 | 514.840 | 139.268 | 33.277 | 22.817 | 68,57% | 6,10 |
| 44 | 483.812 | 131.442 | 33.341 | 22.847 | 68,53% | 5,75 |
| 45 | 597.054 | 160.095 | 33.971 | 24.449 | 71,97% | 6,55 |
| 46 | 581.397 | 154.758 | 34.137 | 24.163 | 70,78% | 6,40 |
| 47 | 568.738 | 158.770 | 34.274 | 24.422 | 71,26% | 6,50 |
| 48 | 496.926 | 138.399 | 34.108 | 23.108 | 67,75% | 5,99 |
| 49 | 541.567 | 156.191 | 34.707 | 24.841 | 71,57% | 6,29 |
| 50 | 504.333 | 149.507 | 34.531 | 24.299 | 70,37% | 6,15 |
| 51 | 427.897 | 121.120 | 34.074 | 21.697 | 63,68% | 5,58 |
| 52 | 434.828 | 131.039 | 34.089 | 22.015 | 64,58% | 5,95 |
| 53 | 439.987 | 131.157 | 34.301 | 23.519 | 68,57% | 5,58 |
| 54 | 565.744 | 167.668 | 34.977 | 25.664 | 73,37% | 6,53 |
| 55 | 517.882 | 157.548 | 35.088 | 25.892 | 73,79% | 6,08 |
| 56 | 609.995 | 182.054 | 35.085 | 26.791 | 76,36% | 6,80 |
| 57 | 538.221 | 167.370 | 34.920 | 26.759 | 76,63% | 6,25 |
| 58 | 593.310 | 184.698 | 35.018 | 28.281 | 80,76% | 6,53 |
| 59 | 613.634 | 188.043 | 35.608 | 29.843 | 83,81% | 6,30 |
| 60 | 407.593 | 121.357 | 32.972 | 24.063 | 72,98% | 5,04 |
| 61 | 658.164 | 209.742 | 32.187 | 32.187 | 100,00% | 6,52 |

[Volver al capítulo: Análisis descriptivo].

2. Apéndice B: Estadísticas de reportes por semana

Tabla 74: Estadísticas descriptivas de reportes y reportadores, semanalmente.

| Semana | Usuarios consultores | Usuarios reportadores | Reportes realizados | Promedio de reportes por usuario | Proporción de usuarios reportadores |
|--------|----------------------|-----------------------|---------------------|----------------------------------|-------------------------------------|
| 1 | 7.200 | 3.047 | 621 | 4,91 | 8,63% |
| 2 | 8.391 | 4.466 | 884 | 5,05 | 10,54% |
| 3 | 8.233 | 4.593 | 841 | 5,46 | 10,21% |
| 4 | 8.573 | 4.120 | 852 | 4,84 | 9,94% |
| 5 | 8.547 | 3.300 | 761 | 4,34 | 8,90% |
| 6 | 9.946 | 4.928 | 1.050 | 4,69 | 10,56% |
| 7 | 10.967 | 5.401 | 1.247 | 4,33 | 11,37% |
| 8 | 8.544 | 3.346 | 782 | 4,28 | 9,15% |
| 9 | 10.768 | 5.976 | 1.136 | 5,26 | 10,55% |
| 10 | 11.778 | 6.073 | 1.186 | 5,12 | 10,07% |
| 11 | 12.041 | 7.037 | 1.257 | 5,60 | 10,44% |
| 12 | 12.368 | 6.123 | 1.299 | 4,71 | 10,50% |
| 13 | 12.939 | 7.756 | 1.545 | 5,02 | 11,94% |
| 14 | 11.829 | 4.593 | 1.072 | 4,28 | 9,06% |
| 15 | 14.076 | 8.242 | 1.691 | 4,87 | 12,01% |
| 16 | 14.684 | 8.648 | 1.660 | 5,21 | 11,30% |
| 17 | 14.879 | 9.065 | 1.743 | 5,20 | 11,71% |
| 18 | 14.414 | 6.763 | 1.402 | 4,82 | 9,73% |
| 19 | 15.448 | 7.849 | 1.711 | 4,59 | 11,08% |
| 20 | 15.717 | 8.126 | 1.623 | 5,01 | 10,33% |
| 21 | 16.330 | 7.995 | 1.725 | 4,63 | 10,56% |
| 22 | 15.622 | 6.572 | 1.545 | 4,25 | 9,89% |
| 23 | 14.849 | 5.862 | 1.380 | 4,25 | 9,29% |
| 24 | 16.563 | 8.545 | 1.718 | 4,97 | 10,37% |
| 25 | 16.908 | 8.869 | 1.795 | 4,94 | 10,62% |
| 26 | 16.875 | 8.414 | 1.814 | 4,64 | 10,75% |
| 27 | 15.927 | 6.618 | 1.451 | 4,56 | 9,11% |
| 28 | 16.136 | 7.348 | 1.721 | 4,27 | 10,67% |
| 29 | 15.726 | 8.616 | 1.679 | 5,13 | 10,68% |
| 30 | 16.005 | 8.322 | 1.631 | 5,10 | 10,19% |
| 31 | 16.735 | 7.329 | 1.455 | 5,04 | 8,69% |
| 32 | 19.935 | 9.744 | 2.053 | 4,75 | 10,30% |
| 33 | 21.887 | 10.264 | 2.331 | 4,40 | 10,65% |
| 34 | 22.665 | 11.080 | 2.415 | 4,59 | 10,66% |

| | | | | | |
|----|--------|--------|-------|------|--------|
| 35 | 22.364 | 10.076 | 2.213 | 4,55 | 9,90% |
| 36 | 24.041 | 11.619 | 2.431 | 4,78 | 10,11% |
| 37 | 23.884 | 10.876 | 2.287 | 4,76 | 9,58% |
| 38 | 23.242 | 8.546 | 2.097 | 4,08 | 9,02% |
| 39 | 24.015 | 9.423 | 2.201 | 4,28 | 9,17% |
| 40 | 22.502 | 7.398 | 1.779 | 4,16 | 7,91% |
| 41 | 24.373 | 9.857 | 2.151 | 4,58 | 8,83% |
| 42 | 24.189 | 9.623 | 2.112 | 4,56 | 8,73% |
| 43 | 22.817 | 7.796 | 1.781 | 4,38 | 7,81% |
| 44 | 22.847 | 7.245 | 1.694 | 4,28 | 7,41% |
| 45 | 24.449 | 8.400 | 1.914 | 4,39 | 7,83% |
| 46 | 24.163 | 7.702 | 1.749 | 4,40 | 7,24% |
| 47 | 24.422 | 7.334 | 1.754 | 4,18 | 7,18% |
| 48 | 23.108 | 6.167 | 1.458 | 4,23 | 6,31% |
| 49 | 24.841 | 6.536 | 1.666 | 3,92 | 6,71% |
| 50 | 24.299 | 6.205 | 1.553 | 4,00 | 6,39% |
| 51 | 21.697 | 5.316 | 1.314 | 4,05 | 6,06% |
| 52 | 22.015 | 5.855 | 1.374 | 4,26 | 6,24% |
| 53 | 23.519 | 5.322 | 1.337 | 3,98 | 5,68% |
| 54 | 25.664 | 7.128 | 1.607 | 4,44 | 6,26% |
| 55 | 25.892 | 6.986 | 1.639 | 4,26 | 6,33% |
| 56 | 26.791 | 7.888 | 1.728 | 4,56 | 6,45% |
| 57 | 26.759 | 7.488 | 1.700 | 4,40 | 6,35% |
| 58 | 28.281 | 8.362 | 1.720 | 4,86 | 6,08% |
| 59 | 29.843 | 8.931 | 1.898 | 4,71 | 6,36% |
| 60 | 24.063 | 5.173 | 1.352 | 3,83 | 5,62% |
| 61 | 32.187 | 8.801 | 2.242 | 3,93 | 6,97% |

[Volver al capítulo: Análisis descriptivo]

3. Apéndice C: Estadísticas de reportes acerca de buses o paradero, por semana.

Tabla 75: Estadísticas semanales de la cantidad de reportes de buses y paraderos, y su proporción.

| Semana | Reportes realizados | Reportes de paraderos | Proporción de paraderos | Reporte de buses | Proporción de buses |
|--------|---------------------|-----------------------|-------------------------|------------------|---------------------|
| 1 | 3.047 | 1.440 | 47,26% | 1.607 | 52,74% |
| 2 | 4.466 | 2.252 | 50,43% | 2.214 | 49,57% |
| 3 | 4.593 | 2.218 | 48,29% | 2.375 | 51,71% |
| 4 | 4.120 | 2.030 | 49,27% | 2.090 | 50,73% |
| 5 | 3.300 | 1.676 | 50,79% | 1.624 | 49,21% |
| 6 | 4.928 | 2.531 | 51,36% | 2.397 | 48,64% |
| 7 | 5.401 | 2.676 | 49,55% | 2.725 | 50,45% |
| 8 | 3.346 | 1.661 | 49,64% | 1.685 | 50,36% |
| 9 | 5.976 | 2.827 | 47,31% | 3.149 | 52,69% |
| 10 | 6.073 | 2.749 | 45,27% | 3.324 | 54,73% |
| 11 | 7.037 | 3.143 | 44,66% | 3.894 | 55,34% |
| 12 | 6.123 | 3.079 | 50,29% | 3.044 | 49,71% |
| 13 | 7.756 | 3.998 | 51,55% | 3.758 | 48,45% |
| 14 | 4.593 | 2.374 | 51,69% | 2.219 | 48,31% |
| 15 | 8.242 | 4.268 | 51,78% | 3.974 | 48,22% |
| 16 | 8.648 | 4.405 | 50,94% | 4.243 | 49,06% |
| 17 | 9.065 | 4.759 | 52,50% | 4.306 | 47,50% |
| 18 | 6.763 | 3.353 | 49,58% | 3.410 | 50,42% |
| 19 | 7.849 | 4.123 | 52,53% | 3.726 | 47,47% |
| 20 | 8.126 | 3.803 | 46,80% | 4.323 | 53,20% |
| 21 | 7.995 | 4.142 | 51,81% | 3.853 | 48,19% |
| 22 | 6.572 | 3.338 | 50,79% | 3.234 | 49,21% |
| 23 | 5.862 | 3.202 | 54,62% | 2.660 | 45,38% |
| 24 | 8.545 | 4.350 | 50,91% | 4.195 | 49,09% |
| 25 | 8.869 | 4.480 | 50,51% | 4.389 | 49,49% |
| 26 | 8.414 | 4.503 | 53,52% | 3.911 | 46,48% |
| 27 | 6.618 | 3.469 | 52,42% | 3.149 | 47,58% |
| 28 | 7.348 | 4.046 | 55,06% | 3.302 | 44,94% |
| 29 | 8.616 | 4.615 | 53,56% | 4.001 | 46,44% |
| 30 | 8.322 | 4.704 | 56,52% | 3.618 | 43,48% |
| 31 | 7.329 | 3.513 | 47,93% | 3.816 | 52,07% |
| 32 | 9.744 | 5.695 | 58,45% | 4.049 | 41,55% |
| 33 | 10.264 | 6.247 | 60,86% | 4.017 | 39,14% |
| 34 | 11.080 | 6.467 | 58,37% | 4.613 | 41,63% |

| | | | | | |
|----|--------|-------|--------|-------|--------|
| 35 | 10.076 | 6.007 | 59,62% | 4.069 | 40,38% |
| 36 | 11.619 | 6.808 | 58,59% | 4.811 | 41,41% |
| 37 | 10.876 | 6.412 | 58,96% | 4.464 | 41,04% |
| 38 | 8.546 | 5.138 | 60,12% | 3.408 | 39,88% |
| 39 | 9.423 | 6.084 | 64,57% | 3.339 | 35,43% |
| 40 | 7.398 | 4.873 | 65,87% | 2.525 | 34,13% |
| 41 | 9.857 | 6.428 | 65,21% | 3.429 | 34,79% |
| 42 | 9.623 | 6.063 | 63,01% | 3.560 | 36,99% |
| 43 | 7.796 | 4.886 | 62,67% | 2.910 | 37,33% |
| 44 | 7.245 | 4.706 | 64,96% | 2.539 | 35,04% |
| 45 | 8.400 | 5.552 | 66,10% | 2.848 | 33,90% |
| 46 | 7.702 | 5.180 | 67,26% | 2.522 | 32,74% |
| 47 | 7.334 | 4.772 | 65,07% | 2.562 | 34,93% |
| 48 | 6.167 | 3.948 | 64,02% | 2.219 | 35,98% |
| 49 | 6.536 | 4.462 | 68,27% | 2.074 | 31,73% |
| 50 | 6.205 | 4.059 | 65,41% | 2.146 | 34,59% |
| 51 | 5.316 | 3.541 | 66,61% | 1.775 | 33,39% |
| 52 | 5.855 | 3.831 | 65,43% | 2.024 | 34,57% |
| 53 | 5.322 | 3.534 | 66,40% | 1.788 | 33,60% |
| 54 | 7.128 | 4.656 | 65,32% | 2.472 | 34,68% |
| 55 | 6.986 | 4.550 | 65,13% | 2.436 | 34,87% |
| 56 | 7.888 | 5.305 | 67,25% | 2.583 | 32,75% |
| 57 | 7.488 | 5.203 | 69,48% | 2.285 | 30,52% |
| 58 | 8.362 | 5.501 | 65,79% | 2.861 | 34,21% |
| 59 | 8.931 | 5.876 | 65,79% | 3.055 | 34,21% |
| 60 | 5.173 | 3.712 | 71,76% | 1.461 | 28,24% |
| 61 | 8.801 | 6.421 | 72,96% | 2.380 | 27,04% |

[Volver al capítulo Reportes de paraderos o buses].

4. Apéndice D: Estadísticas de reportes positivos y pesimistas, según semana.

Tabla 76: Estadísticas semanales de la cantidad de reportes positivos y pesimistas.

| Semana | Reportes realizados | Reportes pesimistas | Proporción de pesimistas | Reportes positivos | Proporción de positivos |
|--------|---------------------|---------------------|--------------------------|--------------------|-------------------------|
| 1 | 3.047 | 1.440 | 47,26% | 1.607 | 52,74% |
| 2 | 4.466 | 2.252 | 50,43% | 2.214 | 49,57% |
| 3 | 4.593 | 2.218 | 48,29% | 2.375 | 51,71% |
| 4 | 4.120 | 2.030 | 49,27% | 2.090 | 50,73% |
| 5 | 3.300 | 1.676 | 50,79% | 1.624 | 49,21% |
| 6 | 4.928 | 2.531 | 51,36% | 2.397 | 48,64% |
| 7 | 5.401 | 2.676 | 49,55% | 2.725 | 50,45% |
| 8 | 3.346 | 1.661 | 49,64% | 1.685 | 50,36% |
| 9 | 5.976 | 2.827 | 47,31% | 3.149 | 52,69% |
| 10 | 6.073 | 2.749 | 45,27% | 3.324 | 54,73% |
| 11 | 7.037 | 3.143 | 44,66% | 3.894 | 55,34% |
| 12 | 6.123 | 3.079 | 50,29% | 3.044 | 49,71% |
| 13 | 7.756 | 3.998 | 51,55% | 3.758 | 48,45% |
| 14 | 4.593 | 2.374 | 51,69% | 2.219 | 48,31% |
| 15 | 8.242 | 4.268 | 51,78% | 3.974 | 48,22% |
| 16 | 8.648 | 4.405 | 50,94% | 4.243 | 49,06% |
| 17 | 9.065 | 4.759 | 52,50% | 4.306 | 47,50% |
| 18 | 6.763 | 3.353 | 49,58% | 3.410 | 50,42% |
| 19 | 7.849 | 4.123 | 52,53% | 3.726 | 47,47% |
| 20 | 8.126 | 3.803 | 46,80% | 4.323 | 53,20% |
| 21 | 7.995 | 4.142 | 51,81% | 3.853 | 48,19% |
| 22 | 6.572 | 3.338 | 50,79% | 3.234 | 49,21% |
| 23 | 5.862 | 3.202 | 54,62% | 2.660 | 45,38% |
| 24 | 8.545 | 4.350 | 50,91% | 4.195 | 49,09% |
| 25 | 8.869 | 4.480 | 50,51% | 4.389 | 49,49% |
| 26 | 8.414 | 4.503 | 53,52% | 3.911 | 46,48% |
| 27 | 6.618 | 3.469 | 52,42% | 3.149 | 47,58% |
| 28 | 7.348 | 4.046 | 55,06% | 3.302 | 44,94% |
| 29 | 8.616 | 4.615 | 53,56% | 4.001 | 46,44% |
| 30 | 8.322 | 4.704 | 56,52% | 3.618 | 43,48% |
| 31 | 7.329 | 3.513 | 47,93% | 3.816 | 52,07% |
| 32 | 9.744 | 5.695 | 58,45% | 4.049 | 41,55% |
| 33 | 10.264 | 6.247 | 60,86% | 4.017 | 39,14% |
| 34 | 11.080 | 6.467 | 58,37% | 4.613 | 41,63% |

| | | | | | |
|----|--------|-------|--------|-------|--------|
| 35 | 10.076 | 6.007 | 59,62% | 4.069 | 40,38% |
| 36 | 11.619 | 6.808 | 58,59% | 4.811 | 41,41% |
| 37 | 10.876 | 6.412 | 58,96% | 4.464 | 41,04% |
| 38 | 8.546 | 5.138 | 60,12% | 3.408 | 39,88% |
| 39 | 9.423 | 6.084 | 64,57% | 3.339 | 35,43% |
| 40 | 7.398 | 4.873 | 65,87% | 2.525 | 34,13% |
| 41 | 9.857 | 6.428 | 65,21% | 3.429 | 34,79% |
| 42 | 9.623 | 6.063 | 63,01% | 3.560 | 36,99% |
| 43 | 7.796 | 4.886 | 62,67% | 2.910 | 37,33% |
| 44 | 7.245 | 4.706 | 64,96% | 2.539 | 35,04% |
| 45 | 8.400 | 5.552 | 66,10% | 2.848 | 33,90% |
| 46 | 7.702 | 5.180 | 67,26% | 2.522 | 32,74% |
| 47 | 7.334 | 4.772 | 65,07% | 2.562 | 34,93% |
| 48 | 6.167 | 3.948 | 64,02% | 2.219 | 35,98% |
| 49 | 6.536 | 4.462 | 68,27% | 2.074 | 31,73% |
| 50 | 6.205 | 4.059 | 65,41% | 2.146 | 34,59% |
| 51 | 5.316 | 3.541 | 66,61% | 1.775 | 33,39% |
| 52 | 5.855 | 3.831 | 65,43% | 2.024 | 34,57% |
| 53 | 5.322 | 3.534 | 66,40% | 1.788 | 33,60% |
| 54 | 7.128 | 4.656 | 65,32% | 2.472 | 34,68% |
| 55 | 6.986 | 4.550 | 65,13% | 2.436 | 34,87% |
| 56 | 7.888 | 5.305 | 67,25% | 2.583 | 32,75% |
| 57 | 7.488 | 5.203 | 69,48% | 2.285 | 30,52% |
| 58 | 8.362 | 5.501 | 65,79% | 2.861 | 34,21% |
| 59 | 8.931 | 5.876 | 65,79% | 3.055 | 34,21% |
| 60 | 5.173 | 3.712 | 71,76% | 1.461 | 28,24% |
| 61 | 8.801 | 6.421 | 72,96% | 2.380 | 27,04% |

[Volver al capítulo: Reportes positivos o pesimistas].

5. Apéndice E: Estadísticas de reportes confirmativos y negativos, según semana.

Tabla 77: Estadísticas semanales de la cantidad de reportes confirmativos y negativos, y su proporción.

| Semana | Reportes realizados | Reportes confirmativos | Proporción de confirmativos | Reportes negativos | Proporción de negativos |
|--------|---------------------|------------------------|-----------------------------|--------------------|-------------------------|
| 1 | 3.047 | 2.837 | 93,11% | 210 | 6,89% |
| 2 | 4.466 | 4.132 | 92,52% | 334 | 7,48% |
| 3 | 4.593 | 4.254 | 92,62% | 339 | 7,38% |
| 4 | 4.120 | 3.763 | 91,33% | 357 | 8,67% |
| 5 | 3.300 | 2.995 | 90,76% | 305 | 9,24% |
| 6 | 4.928 | 4.480 | 90,91% | 448 | 9,09% |
| 7 | 5.401 | 4.947 | 91,59% | 454 | 8,41% |
| 8 | 3.346 | 3.079 | 92,02% | 267 | 7,98% |
| 9 | 5.976 | 5.470 | 91,53% | 506 | 8,47% |
| 10 | 6.073 | 5.618 | 92,51% | 455 | 7,49% |
| 11 | 7.037 | 6.499 | 92,35% | 538 | 7,65% |
| 12 | 6.123 | 5.636 | 92,05% | 487 | 7,95% |
| 13 | 7.756 | 7.080 | 91,28% | 676 | 8,72% |
| 14 | 4.593 | 4.141 | 90,16% | 452 | 9,84% |
| 15 | 8.242 | 7.525 | 91,30% | 717 | 8,70% |
| 16 | 8.648 | 8.025 | 92,80% | 623 | 7,20% |
| 17 | 9.065 | 8.294 | 91,49% | 771 | 8,51% |
| 18 | 6.763 | 6.192 | 91,56% | 571 | 8,44% |
| 19 | 7.849 | 7.092 | 90,36% | 757 | 9,64% |
| 20 | 8.126 | 7.442 | 91,58% | 684 | 8,42% |
| 21 | 7.995 | 7.281 | 91,07% | 714 | 8,93% |
| 22 | 6.572 | 5.988 | 91,11% | 584 | 8,89% |
| 23 | 5.862 | 5.308 | 90,55% | 554 | 9,45% |
| 24 | 8.545 | 7.829 | 91,62% | 716 | 8,38% |
| 25 | 8.869 | 8.076 | 91,06% | 793 | 8,94% |
| 26 | 8.414 | 7.677 | 91,24% | 737 | 8,76% |
| 27 | 6.618 | 5.967 | 90,16% | 651 | 9,84% |
| 28 | 7.348 | 6.585 | 89,62% | 763 | 10,38% |
| 29 | 8.616 | 7.629 | 88,54% | 987 | 11,46% |
| 30 | 8.322 | 7.466 | 89,71% | 856 | 10,29% |
| 31 | 7.329 | 6.634 | 90,52% | 695 | 9,48% |
| 32 | 9.744 | 8.742 | 89,72% | 1.002 | 10,28% |
| 33 | 10.264 | 9.221 | 89,84% | 1.043 | 10,16% |
| 34 | 11.080 | 9.895 | 89,31% | 1.185 | 10,69% |

| | | | | | |
|----|--------|--------|--------|-------|--------|
| 35 | 10.076 | 8.889 | 88,22% | 1.187 | 11,78% |
| 36 | 11.619 | 10.312 | 88,75% | 1.307 | 11,25% |
| 37 | 10.876 | 9.524 | 87,57% | 1.352 | 12,43% |
| 38 | 8.546 | 7.464 | 87,34% | 1.082 | 12,66% |
| 39 | 9.423 | 8.201 | 87,03% | 1.222 | 12,97% |
| 40 | 7.398 | 6.516 | 88,08% | 882 | 11,92% |
| 41 | 9.857 | 8.597 | 87,22% | 1.260 | 12,78% |
| 42 | 9.623 | 8.449 | 87,80% | 1.174 | 12,20% |
| 43 | 7.796 | 6.855 | 87,93% | 941 | 12,07% |
| 44 | 7.245 | 6.255 | 86,34% | 990 | 13,66% |
| 45 | 8.400 | 7.143 | 85,04% | 1.257 | 14,96% |
| 46 | 7.702 | 6.496 | 84,34% | 1.206 | 15,66% |
| 47 | 7.334 | 6.232 | 84,97% | 1.102 | 15,03% |
| 48 | 6.167 | 5.326 | 86,36% | 841 | 13,64% |
| 49 | 6.536 | 5.591 | 85,54% | 945 | 14,46% |
| 50 | 6.205 | 5.189 | 83,63% | 1.016 | 16,37% |
| 51 | 5.316 | 4.466 | 84,01% | 850 | 15,99% |
| 52 | 5.855 | 4.856 | 82,94% | 999 | 17,06% |
| 53 | 5.322 | 4.552 | 85,53% | 770 | 14,47% |
| 54 | 7.128 | 5.975 | 83,82% | 1.153 | 16,18% |
| 55 | 6.986 | 5.997 | 85,84% | 989 | 14,16% |
| 56 | 7.888 | 6.640 | 84,18% | 1.248 | 15,82% |
| 57 | 7.488 | 6.334 | 84,59% | 1.154 | 15,41% |
| 58 | 8.362 | 7.215 | 86,28% | 1.147 | 13,72% |
| 59 | 8.931 | 7.517 | 84,17% | 1.414 | 15,83% |
| 60 | 5.173 | 4.287 | 82,87% | 886 | 17,13% |
| 61 | 8.801 | 7.457 | 84,73% | 1.344 | 15,27% |

[Volver al capítulo: Reportes afirmativos o negativos].

6. Apéndice F: Estadísticas de viajes y evaluaciones semanales

Tabla 78: Cantidad de viajes, evaluaciones y usuarios viajeros y evaluadores, semanalmente.

| Semana | Viajes realizados | Usuarios que viajan | Promedio viajes por usuario | Evaluaciones realizadas | Usuarios que evalúan | Promedio evaluaciones por usuario | Porcentaje viajes evaluados | Porcentaje usuarios que evalúan |
|--------|-------------------|---------------------|-----------------------------|-------------------------|----------------------|-----------------------------------|-----------------------------|---------------------------------|
| 1 | 777 | 360 | 2,16 | 579 | 270 | 2,14 | 74,52% | 75,00% |
| 2 | 1.021 | 506 | 2,02 | 770 | 251 | 3,07 | 75,42% | 49,60% |
| 3 | 903 | 452 | 2,00 | 681 | 188 | 3,62 | 75,42% | 41,59% |
| 4 | 1.037 | 489 | 2,12 | 766 | 206 | 3,72 | 73,87% | 42,13% |
| 5 | 803 | 443 | 1,81 | 614 | 149 | 4,12 | 76,46% | 33,63% |
| 6 | 1.180 | 620 | 1,90 | 828 | 256 | 3,23 | 70,17% | 41,29% |
| 7 | 1.260 | 686 | 1,84 | 928 | 286 | 3,24 | 73,65% | 41,69% |
| 8 | 818 | 477 | 1,71 | 621 | 211 | 2,94 | 75,92% | 44,23% |
| 9 | 1.414 | 702 | 2,01 | 1.071 | 276 | 3,88 | 75,74% | 39,32% |
| 10 | 1.504 | 743 | 2,02 | 1.114 | 272 | 4,10 | 74,07% | 36,61% |
| 11 | 1.574 | 757 | 2,08 | 1.195 | 260 | 4,60 | 75,92% | 34,35% |
| 12 | 1.407 | 732 | 1,92 | 1.042 | 264 | 3,95 | 74,06% | 36,07% |
| 13 | 1.649 | 806 | 2,05 | 1.234 | 250 | 4,94 | 74,83% | 31,02% |
| 14 | 1.023 | 610 | 1,68 | 754 | 202 | 3,73 | 73,70% | 33,11% |
| 15 | 1.748 | 884 | 1,98 | 1.291 | 307 | 4,21 | 73,86% | 34,73% |
| 16 | 1.885 | 914 | 2,06 | 1.356 | 308 | 4,40 | 71,94% | 33,70% |
| 17 | 1.880 | 922 | 2,04 | 1.421 | 309 | 4,60 | 75,59% | 33,51% |
| 18 | 1.487 | 805 | 1,85 | 1.049 | 222 | 4,73 | 70,54% | 27,58% |
| 19 | 1.751 | 895 | 1,96 | 1.272 | 279 | 4,56 | 72,64% | 31,17% |
| 20 | 1.789 | 917 | 1,95 | 1.270 | 292 | 4,35 | 70,99% | 31,84% |
| 21 | 1.763 | 908 | 1,94 | 1.252 | 271 | 4,62 | 71,02% | 29,85% |
| 22 | 1.461 | 795 | 1,84 | 1.053 | 250 | 4,21 | 72,07% | 31,45% |

| | | | | | | | | |
|----|-------|-------|------|-------|-----|------|--------|--------|
| 23 | 1.271 | 723 | 1,76 | 930 | 222 | 4,19 | 73,17% | 30,71% |
| 24 | 1.868 | 908 | 2,06 | 1.353 | 286 | 4,73 | 72,43% | 31,50% |
| 25 | 1.952 | 929 | 2,10 | 1.436 | 301 | 4,77 | 73,57% | 32,40% |
| 26 | 1.874 | 889 | 2,11 | 1.422 | 311 | 4,57 | 75,88% | 34,98% |
| 27 | 1.535 | 719 | 2,13 | 1.119 | 209 | 5,35 | 72,90% | 29,07% |
| 28 | 1.823 | 840 | 2,17 | 1.312 | 245 | 5,36 | 71,97% | 29,17% |
| 29 | 1.555 | 819 | 1,90 | 1.160 | 234 | 4,96 | 74,60% | 28,57% |
| 30 | 1.529 | 777 | 1,97 | 1.173 | 234 | 5,01 | 76,72% | 30,12% |
| 31 | 1.378 | 725 | 1,90 | 1.023 | 172 | 5,95 | 74,24% | 23,72% |
| 32 | 1.846 | 949 | 1,95 | 1.317 | 297 | 4,43 | 71,34% | 31,30% |
| 33 | 2.068 | 1.093 | 1,89 | 1.431 | 331 | 4,32 | 69,20% | 30,28% |
| 34 | 2.194 | 1.106 | 1,98 | 1.545 | 301 | 5,13 | 70,42% | 27,22% |
| 35 | 1.969 | 949 | 2,07 | 1.385 | 253 | 5,47 | 70,34% | 26,66% |
| 36 | 2.167 | 1.073 | 2,02 | 1.553 | 302 | 5,14 | 71,67% | 28,15% |
| 37 | 2.087 | 1.076 | 1,94 | 1.515 | 316 | 4,79 | 72,59% | 29,37% |
| 38 | 1.618 | 899 | 1,80 | 1.158 | 231 | 5,01 | 71,57% | 25,70% |
| 39 | 1.766 | 942 | 1,87 | 1.235 | 221 | 5,59 | 69,93% | 23,46% |
| 40 | 1.270 | 767 | 1,66 | 887 | 214 | 4,14 | 69,84% | 27,90% |
| 41 | 1.832 | 952 | 1,92 | 1.317 | 254 | 5,19 | 71,89% | 26,68% |
| 42 | 1.741 | 884 | 1,97 | 1.183 | 222 | 5,33 | 67,95% | 25,11% |
| 43 | 1.592 | 817 | 1,95 | 1.009 | 149 | 6,77 | 63,38% | 18,24% |
| 44 | 1.539 | 781 | 1,97 | 750 | 128 | 5,86 | 48,73% | 16,39% |
| 45 | 1.804 | 869 | 2,08 | 817 | 146 | 5,60 | 45,29% | 16,80% |
| 46 | 1.468 | 817 | 1,80 | 686 | 114 | 6,02 | 46,73% | 13,95% |
| 47 | 1.482 | 792 | 1,87 | 943 | 172 | 5,48 | 63,63% | 21,72% |
| 48 | 1.219 | 688 | 1,77 | 858 | 147 | 5,84 | 70,39% | 21,37% |
| 49 | 1.314 | 739 | 1,78 | 935 | 163 | 5,74 | 71,16% | 22,06% |
| 51 | 1.323 | 741 | 1,79 | 949 | 164 | 5,79 | 71,73% | 22,13% |
| 52 | 1.125 | 651 | 1,73 | 788 | 136 | 5,79 | 70,04% | 20,89% |

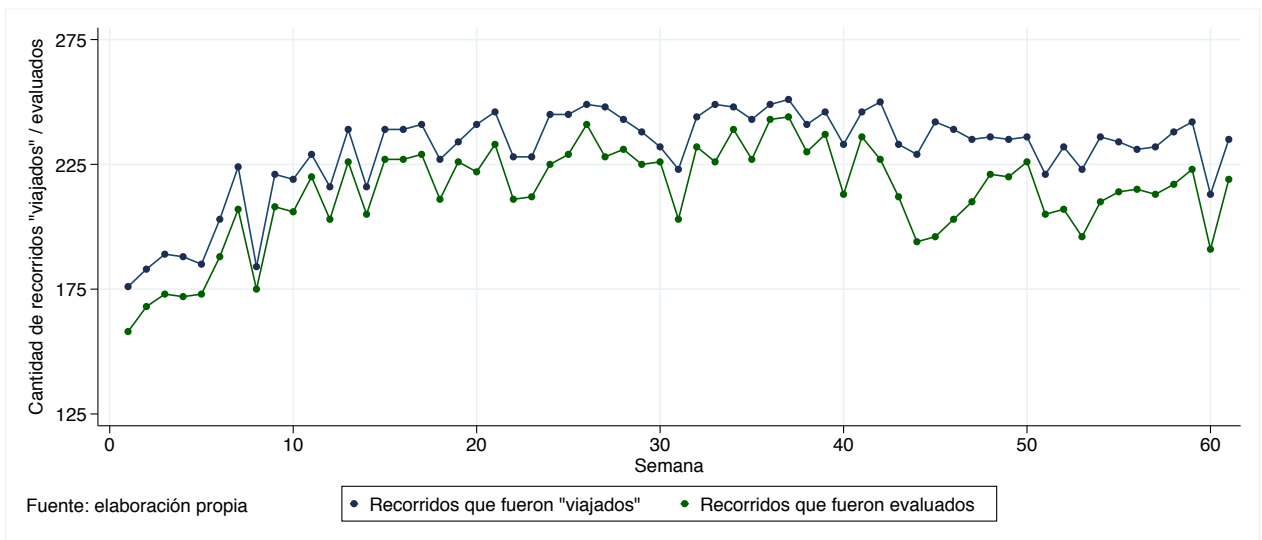
| | | | | | | | | |
|----|-------|-----|------|-------|-----|------|--------|--------|
| 53 | 1.299 | 733 | 1,77 | 899 | 160 | 5,62 | 69,21% | 21,83% |
| 54 | 1.189 | 658 | 1,81 | 814 | 155 | 5,25 | 68,46% | 23,56% |
| 55 | 1.443 | 784 | 1,84 | 969 | 166 | 5,84 | 67,15% | 21,17% |
| 56 | 1.394 | 808 | 1,73 | 958 | 190 | 5,04 | 68,72% | 23,51% |
| 57 | 1.511 | 817 | 1,85 | 1.042 | 187 | 5,57 | 68,96% | 22,89% |
| 58 | 1.277 | 703 | 1,82 | 882 | 159 | 5,55 | 69,07% | 22,62% |
| 59 | 1.506 | 773 | 1,95 | 1.086 | 187 | 5,81 | 72,11% | 24,19% |
| 60 | 1.498 | 805 | 1,86 | 1.048 | 193 | 5,43 | 69,96% | 23,98% |
| 61 | 956 | 583 | 1,64 | 696 | 138 | 5,04 | 72,80% | 23,67% |

[Volver al capítulo: Estadísticas básicas de viajes y evaluaciones].

7. Apéndice G: Estadísticas de recorridos

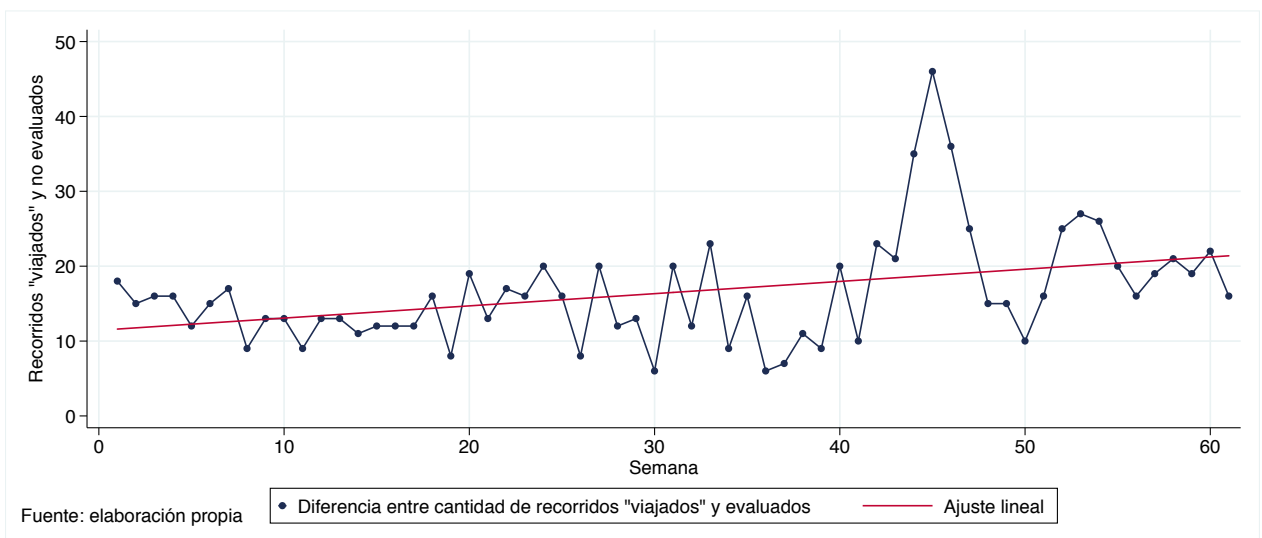
A partir de la Ilustración 23 e Ilustración 24 se concluye que si bien la proporción de usuarios evaluadores disminuye en el tiempo, quienes siguen haciéndolo, evalúan cada vez más. En esta sección se corrobora si esto provoca que exista más cantidad de recorridos viajados pero no evaluados. Para esto, se gráfica la cantidad de recorridos de buses en los cuales al menos un usuario viajó y la cantidad de recorridos en los cuales al menos un usuario evaluó su viaje (ver Ilustración 41). Luego, el espacio entre ambos puntos corresponde a la cantidad de recorridos viajados pero no evaluados.

Ilustración 41: Recorridos semanales de buses que han sido viajados y evaluados por los usuarios.



En la Ilustración 42 se grafica esta diferencia, en donde se le aplica un ajuste lineal.

Ilustración 42: Cantidad de recorridos viajados, pero no evaluados, semanalmente.



Dado lo anterior se desea saber si la cantidad de recorridos no evaluados aumenta con el tiempo, dada la situación planteada anteriormente, para lo cual se propone el siguiente modelo autoregresivo, cuyos resultados se muestran en la Tabla 79. El análisis en el cual se sustenta el siguiente modelo, se encuentra en el siguiente Anexo, Apéndice H: Modelo autoregresivo de los recorridos viajados y no evaluados.

Ecuación 10: Modelo autoregresivo de la cantidad de recorridos viajados y no evaluados, semanalmente.

$$\text{Recorridos no evaluados}_t = B_0 + \sum_{i=1}^4 B_i * \text{Recorridos no evaluados}_{t-i} + B_3 * t + \varepsilon$$

Tabla 79: Resultados estimación modelo autoregresivo de cuatro lags de la cantidad de recorridos viajados y no evaluados, semanalmente.

| Coefficiente | Estimador | Error estándar | Valor t | Pr(> t) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| Semana | 0,127 | 0,054 | 2,32 | 0,024 |
| Rec. no ev. (t-1) | 0,409 | 0,139 | 2,93 | 0,005 |
| Rec. no ev. (t-2) | 0,382 | 0,143 | 2,67 | 0,010 |
| Rec. no ev. (t-3) | -0,358 | 0,143 | -2,50 | 0,016 |
| Rec. no ev. (t-4) | -0,009 | 0,138 | -0,67 | 0,504 |
| Intercepto | 6,653 | 2,356 | 2,82 | 0,007 |

Se tiene que la serie de tiempo, presenta una tendencia positiva significativa al 5%, significando que por cada semana que transcurre aumenta en 0,127 los recorridos no evaluados pero que sí fueron evaluados. Por lo que se verifica la idea anterior, cada vez hay menos evaluadores y quienes lo hacen, lo hacen más, provocando que haya menos recorridos evaluados a medida que transcurre el tiempo.

[Volver al capítulo: Estadísticas básicas de viajes y evaluaciones].

8. Apéndice H: Modelo autoregresivo de los recorridos viajados y no evaluados

Para analizar la cantidad de recorridos no evaluados en el tiempo, primero se propone un modelo simple, con la siguiente especificación.

Ecuación 11: Modelo simple de la cantidad de recorridos viajados y no evaluados, semanalmente.

$$\text{Recorridos no evaluados}_t = B_0 + B_1 * t + \varepsilon$$

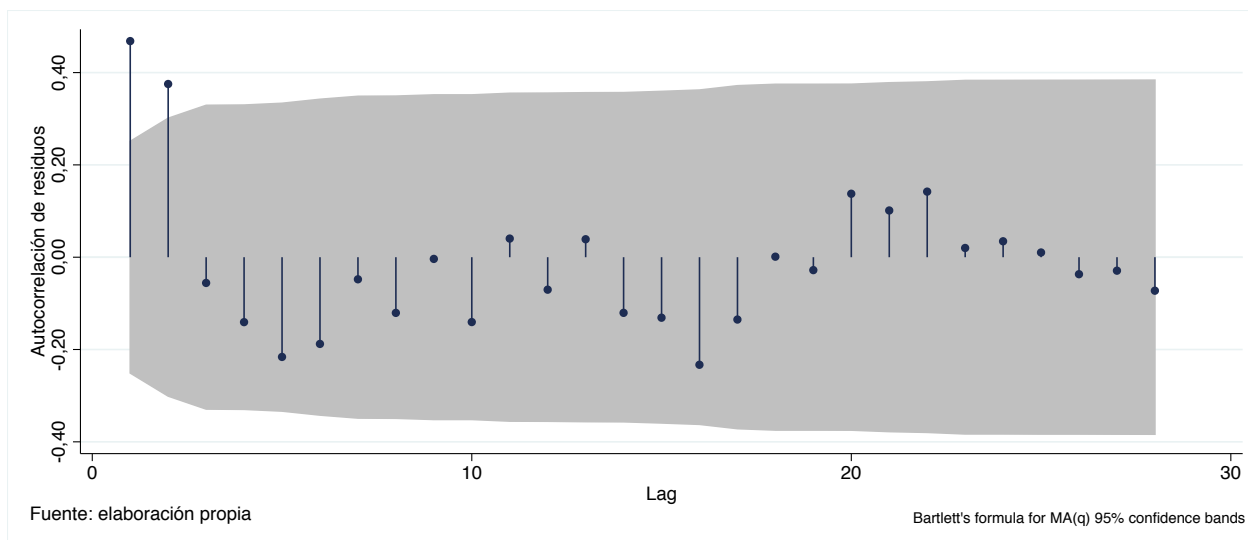
Los resultados de la estimación se encuentran en la Tabla 80.

Tabla 80: Resultados estimación modelo simple de recorridos no evaluados.

| Coefficiente | Estimador | Error estándar | Valor t | Pr(> t) |
|--------------|-----------|----------------|---------|----------|
| Semana (t) | 0,163 | 0,049 | 3,31 | 0,002 |
| Intercepto | 11,434 | 1,755 | 6,51 | 0,000 |

Los residuos de esta estimación se grafican, en donde se observa una autocorrelación de los residuos 1 y 2 (ver Ilustración 43).

Ilustración 43: Autocorrelación de los residuos del modelo simple.



Para confirmar lo anterior, se realiza un test de Breusch-Godfrey para identificar autocorrelación de los errores, cuyos resultados indican que sí la hay (ver Tabla 81).

Tabla 81: Resultados test Breusch-Godfrey para modelo simple.

| Lags (ρ) | Estadístico χ^2 | Pr(> χ^2) |
|-----------------|----------------------|-----------------|
| 1 | 13,523 | 0,0002 |
| 2 | 15,388 | 0,0005 |
| 3 | 22,361 | 0,0001 |
| 4 | 22,715 | 0,0001 |

Dado estos resultados, se propone un modelo autoregresivo de dos rezagos, cuya especificación y resultados se encuentran en la Ecuación 12 y Tabla 82, respectivamente.

Ecuación 12: Modelo autoregresivo de la cantidad de recorridos viajados y no evaluados, semanalmente.

$$\text{Recorridos no evaluados}_t = B_0 + \sum_{i=1}^2 B_i * \text{Recorridos no evaluados}_{t-i} + B_3 * t + \varepsilon$$

Tabla 82: Resultados modelo autoregresivo de dos lags.

| Coefficiente | Estimador | Error estándar | Valor t | Pr(> t) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| Semana (t) | 0,077 | 0,051 | 1,49 | 0,143 |
| Rec. no ev. (t-1) | 0,374 | 0,133 | 2,81 | 0,007 |
| Rec. no ev. (t-2) | 0,201 | 0,132 | 1,52 | 0,133 |
| Intercepto | 4,552 | 2,221 | 2,06 | 0,044 |

Se realiza nuevamente el test de Breusch-Godfrey en donde se rechaza la hipótesis nula (ver Tabla 83), lo que significa la presencia de autocorrelación.

Tabla 83: Resultados test Breusch-Godfrey para modelo autoregresivo de dos lags.

| Lags (ρ) | Estadístico χ^2 | Pr(>χ^2) |
|---------------------------------|--|------------------------------------|
| 1 | 8,939 | 0,0028 |
| 2 | 9,049 | 0,0108 |
| 3 | 9,185 | 0,0269 |
| 4 | 9,509 | 0,0496 |

Dado lo anterior, se extiende los rezagos a 4 períodos, cuyos resultados de la estimación están en la Tabla 84.

Tabla 84: Resultados estimación modelo autoregresivo de cuatro lags de la cantidad de recorridos viajados y no evaluados, semanalmente.

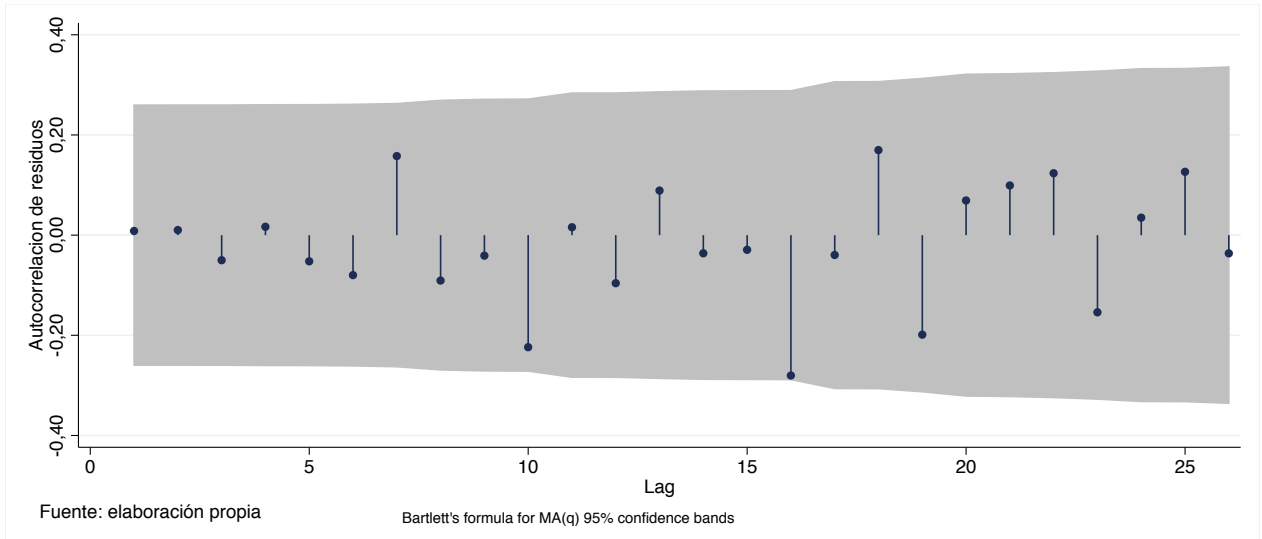
| Coefficiente | Estimador | Error estándar | Valor t | Pr(> t) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| Semana (t) | 0,127 | 0,054 | 2,32 | 0,024 |
| Rec. no ev. (t-1) | 0,409 | 0,139 | 2,93 | 0,005 |
| Rec. no ev. (t-2) | 0,382 | 0,143 | 2,67 | 0,010 |
| Rec. no ev. (t-3) | -0,358 | 0,143 | -2,50 | 0,016 |
| Rec. no ev. (t-4) | -0,009 | 0,138 | -0,67 | 0,504 |
| Intercepto | 6,653 | 2,356 | 2,82 | 0,007 |

El test de Breusch-Godfrey indica que ya no hay presencia de autocorrelación (ver Tabla 85), y al graficar los residuos del modelo autoregresivo de 4 lags tampoco se observa autocorrelación (ver Ilustración 44).

Tabla 85: Resultados test Breusch-Godfrey del modelo autoregresivo de 4 lags.

| Lags (ρ) | Estadístico χ^2 | Pr(>χ^2) |
|---------------------------------|--|------------------------------------|
| 1 | 0,208 | 0,6482 |
| 2 | 0,246 | 0,8840 |
| 3 | 0,807 | 0,8477 |
| 4 | 0,982 | 0,9126 |

Ilustración 44: Residuos de modelo autoregresivo de 4 lags.



9. Apéndice I: Notas de la base de datos

Los valores utilizados en las evaluaciones se indican en la Tabla 86.

Tabla 86: Valores de evaluaciones en el informe según base de datos.

| Nota para análisis informe | Nota descrita en la base de datos |
|----------------------------|-----------------------------------|
| 0 | 0 |
| 1 | 1, 11 |
| 2 | 2, 12 |
| 3 | 3, 4, 13 |
| 4 | 5, 14 |
| 5 | 7, 15 |

[Volver al capítulo: Análisis descriptivo de la nota del viaje].

10. Apéndice J: Velocidad del viaje según día de la semana e intervalo horario

En el capítulo 3.4C, subcapítulo “Variación local de la evaluación, según duración y tiempo del viaje” se muestra que la evaluación del recorrido presenta variaciones según la velocidad del bus. Es conocido que la velocidad del bus varía dependiendo del día y la hora en la ciudad de Santiago (DTPM, 2019), por lo que se busca identificar si la velocidad presenta variaciones según estas variables utilizando los datos recopilados por los usuarios de la aplicación. Notar que en este capítulo se transforma la variable metros/minuto a kilómetros/hora para facilitar el entendimiento.

Ecuación 13: Especificación modelo para explicar la velocidad del bus según el intervalo horario.

$$Velocidad\ bus_i = B_o + \sum_{j=1}^6 B_j * Intervalo\ horario_{ji} + \varepsilon$$

En donde los intervalos horarios son de 3 horas a partir de la hora punta mañana, a excepción de la noche-madrugada que se extiende a 6 horas, el cual además es el grupo base en esta regresión. Los resultados de ésta se encuentran en la Tabla 87.

Tabla 87: Resultados regresión velocidad del bus según intervalo horario.

| Coefficiente | Estimador | Error estándar | Valor t | Pr(> t) |
|---------------|-----------|----------------|---------|----------|
| 6:01 a 9:00 | -5,432 | 0,244 | -22,23 | 0,000 |
| 9:01 a 12:00 | -4,860 | 0,249 | -19,45 | 0,000 |
| 12:01 a 15:00 | -5,358 | 0,248 | -21,56 | 0,000 |
| 15:01 a 18:00 | -5,520 | 0,247 | -22,31 | 0,000 |
| 18:01 a 21:00 | -6,410 | 0,245 | -26,08 | 0,000 |
| 21:01 a 00:00 | -3,191 | 0,245 | -12,53 | 0,000 |
| Intercepto | 19,250 | 0,233 | 82,48 | 0,000 |

Se concluye que la velocidad de los buses es más alta en el grupo de base, durante las 00:01 y 6:00, y que la diferencia de ésta con respecto a los otros intervalos es significativa al 1%.

En la Tabla 88 se realiza una regresión similar, pero con variables independientes los días de las semanas, su especificación es la siguiente. Notar que en esta oportunidad el grupo base es el domingo.

Ecuación 14: Especificación modelo para explicar la velocidad del bus según el día de la semana.

$$Velocidad\ bus_i = B_o + \sum_{j=1}^6 B_j * Día\ Semana_{ji} + \varepsilon$$

Los resultados de este modelo están en la Tabla 88, en donde se observa que el día con mayor velocidad es el domingo, cuya diferencia es significativa al 1% respecto a los otros días, que tienen velocidades menores en torno a los 2,5 kilómetros por hora.

Tabla 88: Resultados regresión velocidad del bus según día de la semana.

| Coefficiente | Estimador | Error estándar | Valor t | Pr(> t) |
|---------------------|------------------|-----------------------|----------------|--------------------|
| Lunes | -2,656 | 0,147 | -18,00 | 0,000 |
| Martes | -2,271 | 0,147 | -15,44 | 0,000 |
| Miércoles | -2,517 | 0,146 | -17,22 | 0,000 |
| Jueves | -2,567 | 0,147 | -17,47 | 0,000 |
| Viernes | -2,537 | 0,147 | -17,15 | 0,000 |
| Sábado | -1,086 | 0,158 | -6,86 | 0,000 |
| Intercepto | 16,225 | 0,121 | 134,09 | 0,000 |

De ambas tablas construye la Tabla 89, en donde se especifica la velocidad estimada de cada intervalo horario y día de la semana.

Tabla 89: Resultados generales de estimación de velocidad de bus según intervalo horario y día de la semana.

| Intervalo horario | Velocidad [km/hr] | | Día de la semana | Velocidad [km/hr] |
|--------------------------|--------------------------|--|-------------------------|--------------------------|
| 6:01 a 9:00 | 13,82 | | Lunes | 13,57 |
| 9:01 a 12:00 | 14,39 | | Martes | 13,95 |
| 12:01 a 15:00 | 13,89 | | Miércoles | 13,71 |
| 15:01 a 18:00 | 13,73 | | Jueves | 13,66 |
| 18:01 a 21:00 | 12,84 | | Viernes | 13,69 |
| 21:01 a 00:00 | 16,06 | | Sábado | 15,14 |
| 00:01 a 6:00 | 19,25 | | Domingo | 16,23 |

[Volver al capítulo: ¡Error! No se encuentra el origen de la referencia.]

11. Apéndice K: Condiciones experimentales

Tabla 90: Configuración y detalle de la notificación de cada condición experimental del día 2.

| Apartado | Sorteo-regular | Sorteo-variante | Prosocial | Ambas |
|---|--|---|---|---|
| Título de notificación | Reporta, no te quedes sin participar por cargas bip! | Reporta, no te quedes sin participar por cargas bip! | Reporta, no te quedes sin ayudar | Reporta, no te quedes sin ayudar y participar por cargas bip! |
| Subtítulo de notificación | Revisa tu bandeja de entrada para leer más 🐸 | Revisa tu bandeja de entrada para leer más 🐸 | Revisa tu bandeja de entrada para leer más 🐸 | Revisa tu bandeja de entrada para leer más 🐸 |
| Mensaje particular de cada condición experimental | Reporta sobre un paradero hasta mañana y así estarás participando en un sorteo por tres cargas bip! de \$10.000. | Reporta sobre un paradero hasta mañana y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de \$10.000. | Reporta sobre un paradero hasta mañana y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. | Reporta sobre un paradero hasta mañana y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de \$10.000. |
| Mensaje común para todas las condiciones experimentales | <p>Recuerda que para reportar sobre un paradero cercano a tu ubicación, primero debes presionar el paradero en el mapa, y luego en el botón amarillo que saldrá en la parte inferior de la pantalla.</p> <p>Transapp 🐸</p> | | | |

Tabla 91: Configuración y detalle de la notificación de cada condición experimental del día 3.

| Apartado | Sorteo-regular | Sorteo-variante | Prosocial | Ambas |
|---|--|--|--|--|
| Título de notificación | Último día: Participa por cargas bip! de \$10.000! 🎉 | Último día: Participa por cargas bip! de \$10.000! 🎉 | Ultimo día: Ayuda al resto de los viajeros! 🎉 | Último día: Ayuda al resto y participa por cargas bip! 🎉 |
| Subtítulo de notificación | Campaña: Gana reportando tu paradero | Campaña: Gana reportando tu paradero | Campaña: Ayuda reportando tu paradero | Campaña: Ayuda y gana reportando tu paradero |
| Mensaje particular de cada condición experimental | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y estarás participando en un sorteo por tres cargas bip! de \$10.000. | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de \$10.000. | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. | Participa en esta campaña entre hoy y el jueves 17 de Octubre reportando sobre un paradero, y así estarás ayudando a mejorar el viaje de otras personas y el transporte público. También estarás participando en un sorteo por tres cargas bip! de \$10.000. |
| Mensaje común para todas las condiciones experimentales | <p>Recuerda que para reportar sobre un paradero cercano a tu ubicación, primero debes presionar el paradero en el mapa, y luego en el botón amarillo que saldrá en la parte inferior de la pantalla.</p> <p>Gracias por ser parte de nuestra comunidad 🍀</p> | | | |

[Volver al capítulo: Condiciones experimentales].

12. Apéndice L: Análisis de resultados de experimento considerando usuarios que no reciben las notificaciones

Tal como se indica, hubo un problema en el envío de las notificaciones en la aplicación, lo que provocó que sólo el 37,91% de los usuarios recibiera al menos una de las notificaciones. Las causales son desconocidos para el equipo de la aplicación y se considera que la eliminación de los datos asociados a los usuarios no receptores para el cálculo de los resultados del experimento no es perjudicial, lo cual se comprueba a continuación.

Tabla 92: Proporción de usuarios no receptores de notificaciones que reportaron, según condición experimental.

| Condición experimental | Usuarios que no reciben notif. | Usuarios que reportan | Proporción de usuarios |
|------------------------|--------------------------------|-----------------------|------------------------|
| “Económico” | 6.941 | 58 | 0,8356% |
| “Econ. – txt. Pr.” | 6.895 | 46 | 0,6671% |
| “Prosocial” | 4.073 | 18 | 0,4419% |
| “Ambos” | 6.930 | 40 | 0,5772% |
| “Control” | 6.512 | 44 | 0,6756% |

Se realiza una regresión logit con variable dependiente “Reporta” e independiente las diferentes condiciones experimentales (la utilizada para estimar el “ITT”), con el objetivo de ver si hay diferencias estadísticas significativas en la proporción de usuarios que reportan entre las condiciones experimentales. Los resultados se muestran en la Tabla 93.

Tabla 93: Resultados estimación modelo logit en usuarios no receptores de notificaciones.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) |
|--------------------|-----------|----------------|---------|----------|
| “Económico” | 0,214 | 0,200 | 1,07 | 0,286 |
| “Econ. – txt. Pr.” | -0,013 | 0,212 | -0,06 | 0,952 |
| “Prosocial” | -0,427 | 0,281 | -1,52 | 0,128 |
| “Ambos” | -0,159 | 0,219 | -0,72 | 0,469 |
| Intercepto | -4,990 | 0,151 | -32,99 | 0,000 |

Como ningún coeficiente asociado a las diferentes condiciones experimentales es significativo, se concluye que, para los usuarios que no recibieron notificación, no hay evidencia estadística que compruebe que existen diferencias significativas con respecto el grupo de control, así permitiendo eliminar estos individuos del análisis final del experimento.

[Volver al capítulo: Exposición al tratamiento].

13. Apéndice M: Análisis de resultados omitiendo segundo día

En este apéndice sólo se indica las principales tablas de resultados de los enfoques de recepción y apertura, omitiendo la data de recepción y apertura de las notificaciones del segundo día del experimento. Es decir, sólo se considera la información del primer y tercer día, no obstante, para la variable de interés dependiente, si la persona reporta o no, de igual manera se considera el segundo día ya que esta acción es independiente si las personas recibieron o no otras notificaciones y/o si hay desconocimiento del investigador en cuál recibió y/o abrió.

13.1 Resultados según recepción

En la Tabla 94 se muestran los resultados de la estimación del “ITT” en el modelo – simple. Los cuales no varían en gran magnitud respecto a los mostrados en el capítulo de análisis. En particular, las notificaciones “Económicos” y “Prosocial” en esta estimación difieren en 0,07% en su efecto respecto al mostrado en el informe.

Tabla 94: Resultados modelo logit - simple, para estimación del efecto ITT omitiendo datos del segundo día.

| Coefficiente | Estimador | Error estándar | Valor Z | Pr(> z) | Odds-ratio | Efecto marginal³⁵ |
|---------------------|------------------|-----------------------|----------------|--------------------|-------------------|-------------------------------------|
| “Económicos” | 2,692 | 0,156 | 17,25 | 0,000 | 14,76 | 8,45% |
| “Prosocial” | 1,345 | 0,200 | 6,72 | 0,000 | 3,84 | 1,87% |
| “Ambos” | 2,283 | 0,164 | 13,89 | 0,000 | 9,81 | 5,58% |
| Intercepto | -4,990 | 0,151 | -32,99 | 0,000 | 0,006 | - |

A continuación, los resultados de la estimación del modelo con heterogeneidad del tratamiento según tipo de usuario (ver Tabla 95). Se observa que todas las variables asociadas a notificaciones son significativas al 1% de igual forma que en los resultados principales del trabajo de título.

³⁵ Efecto marginal promedio (*Average marginal effect*, “AME”) respecto al grupo control, medido en probabilidad de emitir reporte. Resultado obtenido utilizando comando “*margins, dydx*” en *Stata*.

Tabla 95: Resultados modelo logit - tipo de usuario, para estimación del efecto ITT, omitiendo datos del segundo día.

| Coeficiente | | Estimador | Error estándar | Valor Z | Pr(> z) | Odds-ratio | Efecto marginal ³⁶ |
|---------------------|--------------|-----------|----------------|---------|----------|------------|-------------------------------|
| "Reportador" | "Económicos" | 5,379 | 0,300 | 17,93 | 0,000 | 216,95 | 35,90% |
| | "Prosocial" | 5,477 | 0,369 | 14,82 | 0,000 | 239,29 | 38,20% |
| | "Ambos" | 5,023 | 0,351 | 14,30 | 0,000 | 152,02 | 28,16% |
| | "Control" | 3,757 | 0,375 | 10,00 | 0,000 | 42,83 | 9,80% |
| "Reportador Tímido" | "Económicos" | 4,256 | 0,278 | 15,30 | 0,000 | 70,51 | 15,29% |
| | "Prosocial" | 2,500 | 0,397 | 6,30 | 0,000 | 12,18 | 2,82% |
| | "Ambos" | 4,088 | 0,292 | 13,98 | 0,000 | 59,62 | 13,21% |
| | "Control" | 1,715 | 0,379 | 4,52 | 0,000 | 5,55 | 1,17% |
| "No Reportador" | "Económicos" | 3,347 | 0,272 | 12,31 | 0,000 | 28,42 | 6,65% |
| | "Prosocial" | 1,543 | 0,340 | 4,54 | 0,000 | 4,68 | 0,95% |
| | "Ambos" | 2,851 | 0,281 | 10,16 | 0,000 | 17,31 | 4,06% |
| Intercepto | | -5,947 | 0,267 | -22,23 | 0,000 | 0,002 | - |

Luego se realiza el cambio de base para facilitar la lectura del efecto de la recepción de la notificación (ver Tabla 96). En donde los resultados presentan una variación levemente mayor, por ejemplo, en los usuarios "Reportadores" hay una diferencia de 0,59% en el grupo "Prosocial", la cual es la más alta entre todas las condiciones experimentales, lo que se podría deber a que la data de usuarios de este grupo se vio disminuida. En el grupo "masivo" de "No Reportadores" la diferencia tiene su máximo valor en 0,11% en el grupo "Prosocial" (el cual es el más reducido dentro de su categoría en términos de cantidad de usuarios).

Tabla 96: Resultados de efecto ITT a partir de modelo - tipo de usuario, omitiendo datos del segundo día.

| Condición experimental | Reportadores | Reportadores tímidos | No Reportadores |
|------------------------|--------------|----------------------|-----------------|
| "Económicos" | 26,10% | 14,12% | 6,65% |
| "Prosocial" | 28,40% | 1,65% | 0,95% |
| "Ambos" | 18,36% | 12,04% | 4,06% |
| "Control" | Base | Base | Base |

Se concluye que al menos en términos de resultados según recepción, no hay mayor diferencia notable en magnitud y significancia de los términos importantes asociados a los efectos de las diferentes notificaciones en la probabilidad de emitir un reporte sobre un paradero.cc

³⁶ Efecto marginal promedio (*Average marginal effect*, AME) respecto al grupo "Control - No Reportador", medido en probabilidad de emitir reporte. Resultado obtenido utilizando comando "*margins, dydx*" en *Stata*.

13.2 Resultados según apertura

En la Tabla 97 se muestra los resultados del modelo 2SLS para la estimación del efecto “CACE” de las diferentes notificaciones controlando por heterogeneidad del tratamiento según tipo de usuario.

Tabla 97: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, omitiendo datos del segundo día.

| Coeficiente | | Estimador | Error estándar | Valor Z | Pr(> z) | Efecto marginal |
|---------------------|--------------------|-----------|----------------|---------|----------|-----------------|
| “Reportador” | “Económico” | 1,0394 | 0,051 | 20,12 | 0,000 | 103,94% |
| | “Econ. – txt. Pr.” | 0,9449 | 0,053 | 17,64 | 0,000 | 94,49% |
| | “Prosocial” | 1,5519 | 0,103 | 15,00 | 0,000 | 155,19% |
| | “Ambos” | 0,9225 | 0,069 | 13,33 | 0,000 | 92,25% |
| | “Control” | 0,0980 | 0,016 | 5,97 | 0,000 | 9,80% |
| “Reportador Tímido” | “Económico” | 0,5527 | 0,030 | 18,18 | 0,000 | 55,27% |
| | “Econ. – txt. Pr.” | 0,6247 | 0,034 | 18,14 | 0,000 | 62,47% |
| | “Prosocial” | 0,3433 | 0,130 | 2,64 | 0,008 | 34,33% |
| | “Ambos” | 0,4963 | 0,032 | 15,22 | 0,000 | 49,63% |
| | “Control” | 0,0117 | 0,007 | 1,65 | 0,099 | 1,17% |
| “No Reportador” | “Económico” | 0,3876 | 0,024 | 16,02 | 0,000 | 38,76% |
| | “Econ. – txt. Pr.” | 0,3299 | 0,024 | 13,59 | 0,000 | 32,99% |
| | “Prosocial” | 0,1179 | 0,067 | 1,74 | 0,082 | 11,79% |
| | “Ambos” | 0,2860 | 0,031 | 9,09 | 0,000 | 28,60% |
| Intercepto | | 0,0026 | 0,003 | 0,94 | 0,349 | - |

En términos de significancia de las variables, destaca la variable “Prosocial” en la categoría de “No Reportador” que aumenta su p-valor de 0,042 a 0,082, provocando que no sea significativa al 5% pero sí al 10%. El resto de las variables mantiene prácticamente idéntico su p-valor. Respecto a la magnitud de los valores, se procede a realizar el cambio de base (ver Tabla 98).

Tabla 98: Resultados de efecto CACE a partir de modelo - simple, omitiendo datos del segundo día.

| Condición experimental | Reportadores | | Reportadores Tímidos | | No Reportadores | |
|------------------------|--------------|---------|----------------------|---------|-----------------|--------|
| | | | | | | |
| “Económico” | 94,14% | +10,70% | 55,27% | +11,03% | 38,76% | +6,76% |
| “Econ. – txt. Pr.” | 84,69% | +19,15% | 62,47% | +9,48% | 32,99% | +5,62% |
| “Prosocial” | 145,39% | +51,74% | 34,33% | +13,02% | 11,79% | +1,65% |
| “Ambos” | 94,14% | +32,09% | 49,63% | +6,24% | 28,60% | +5,06% |
| “Control” | Base | | Base | | Base | |

En esta categoría es donde se presenta mayor cambio de magnitud de los diferentes valores, dado que todas las notificaciones aumentaron su magnitud dada la apertura de la notificación respecto al grupo de control. En la categoría de “Reportadores” la diferencia es máxima en el grupo “Prosocial” con un 52% el cual podría deberse a que en este grupo hay sólo 65 personas receptoras de la notificación de las cuales 16 la abren y 11 reportan, provocando así que ligeros aumentos de cantidad de aperturas y recepciones aumenten en gran medida los porcentajes (a diferencia de los otros grupos donde son más de un centenar de personas en las diferentes condiciones experimentales).

[Volver al capítulo: *Exposición al tratamiento*].

14. Apéndice N: Resultados de regresión lineal 2SLS – para estimación CACE modelo tipo de usuario

En esta sección se muestra los resultados de la estimación del modelo de regresión lineal de 2SLS para el cálculo del CACE en los modelos de tipos de usuario. En el informe se adjunta cuando el grupo base es el control de la clasificación “No Reportador” de los tipos de usuario. En este anexo, se muestra los resultados considerando como base “Reportador Tímido” y “Reportador”.

14.1 Grupo base “Reportador Tímido”

Tabla 99: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, con grupo base control del tipo de usuario “Reportador Tímido”.

| Coeficiente | | Estimador | Error estándar | Valor Z | Pr(> z) | Efecto marginal |
|---------------------|--------------------|-----------|----------------|---------|----------|-----------------|
| “Reportador” | “Económico” | 0,9387 | 0,044 | 21,17 | 0,000 | 93,87% |
| | “Econ. – txt. Pr.” | 0,7592 | 0,041 | 18,64 | 0,000 | 75,92% |
| | “Prosocial” | 1,0416 | 0,067 | 15,45 | 0,000 | 104,16% |
| | “Ambos” | 0,7250 | 0,050 | 14,56 | 0,000 | 72,50% |
| | “Control” | 0,1006 | 0,016 | 6,37 | 0,000 | 10,06% |
| “Reportador Tímido” | “Económico” | 0,4502 | 0,023 | 19,82 | 0,000 | 45,02% |
| | “Econ. – txt. Pr.” | 0,5389 | 0,027 | 20,30 | 0,000 | 53,89% |
| | “Prosocial” | 0,2321 | 0,072 | 3,23 | 0,001 | 23,21% |
| | “Ambos” | 0,4422 | 0,025 | 17,51 | 0,000 | 44,22% |
| “No Reportador” | “Económico” | 0,3315 | 0,015 | 21,92 | 0,000 | 33,15% |
| | “Econ. – txt. Pr.” | 0,2855 | 0,015 | 18,63 | 0,000 | 28,55% |
| | “Prosocial” | 0,1262 | 0,043 | 2,95 | 0,003 | 12,62% |
| | “Ambos” | 0,2508 | 0,020 | 12,57 | 0,000 | 25,08% |
| | “Control” | 0,0026 | 0,003 | 0,96 | 0,338 | 0,26% |

Se desprende que todos los efectos de las aperturas de las diferentes notificaciones son significativos respecto al grupo de control de los “Reportadores Tímidos”.

14.2 Grupo base “Reportador”

Tabla 100: Resultados modelo regresión lineal - tipo de usuario, para estimación del efecto CACE, con grupo base control del tipo de usuario “Reportado”.

| Coeficiente | | Estimador | Error estándar | Valor Z | Pr(> z) | Efecto marginal |
|---------------------|--------------------|-----------|----------------|---------|----------|-----------------|
| “Reportador” | “Económico” | 0,9387 | 0,044 | 21,15 | 0,000 | 93,87% |
| | “Econ. – txt. Pr.” | 0,7592 | 0,041 | 18,62 | 0,000 | 75,92% |
| | “Prosocial” | 1,0416 | 0,067 | 15,44 | 0,000 | 104,16% |
| | “Ambos” | 0,7250 | 0,050 | 14,55 | 0,000 | 72,50% |
| “Reportador Tímido” | “Económico” | 0,4502 | 0,023 | 19,80 | 0,000 | 45,02% |
| | “Econ. – txt. Pr.” | 0,5389 | 0,027 | 20,28 | 0,000 | 53,89% |
| | “Prosocial” | 0,2321 | 0,072 | 3,22 | 0,001 | 23,21% |
| | “Ambos” | 0,4422 | 0,025 | 17,49 | 0,000 | 44,22% |
| | “Control” | 0,0143 | 0,006 | 2,24 | 0,025 | 1,43% |
| “No Reportador” | “Económico” | 0,3315 | 0,015 | 21,91 | 0,000 | 33,15% |
| | “Econ. – txt. Pr.” | 0,2855 | 0,015 | 18,61 | 0,000 | 28,55% |
| | “Prosocial” | 0,1262 | 0,043 | 2,94 | 0,003 | 12,62% |
| | “Ambos” | 0,2508 | 0,020 | 12,55 | 0,000 | 25,08% |
| | “Control” | 0,0026 | 0,003 | 0,96 | 0,338 | 0,26% |

Se desprende que todos los efectos de las aperturas de las diferentes notificaciones son significativos respecto al grupo de control de los “Reportadores”.

[Volver al capítulo: Análisis – modelo según tipo de usuario].